CREATIVE SENSE-MAKING: A COGNITIVE FRAMEWORK FOR QUANTIFYING INTERACTION DYNAMICS IN CO-CREATION

A Thesis
Presented to
The Academic Faculty

by

Nicholas M. Davis

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Interactive Computing

Georgia Institute of Technology
May 2017

CREATIVE SENSE-MAKING: A COGNITIVE FRAMEWORK FOR QUANTIFYING INTERACTION DYNAMICS IN CO-CREATION

Approved by:

Professor Brian Magerko, Advisor
School of Interactive Computing
Georgia Institute of Technology

Professor Ellen Yi-Luen Do,
Co-Advisor
School of Interactive Computing
Georgia Institute of Technology

Professor Ashok Goel
School of Interactive Computing
Georgia Institute of Technology

Professor Mark Riedl
School of Interactive Computing
Georgia Institute of Technology

Professor Michael Nitsche
School of Literature, Media, and
Communication
Georgia Institute of Technology

Professor Mary Lou Maher
College of Computing and Informatics
University of North Carolina Charlotte

Date Approved: November 29, 2016

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Brian Magerko, for his continuous support in the Human-Centered Computing Ph.D. program. My work exploring co-creative drawing found a permanent home in Brians lab a few years after I started at Georgia Tech, situated among other highly related co-creative systems and empirical investigations into creativity. Brians passion for investigating cognition with creative technologies facilitated many discussions about the role of enaction and sense-making in cognitive architectures, critically helping form and refine my thinking about designing and evaluating co-creative systems. I am grateful to all the members of the ADAM lab, especially Mikhail Jacob, whos LuminAI co-creative dance agent provided many opportunities for collaboration and knowledge sharing. The work presented in this thesis would not have been possible without these discussions and support from the ADAM lab. I am also extremely grateful for the understanding and support Brian showed for the health challenges I faced during the program.

I would also like to extend a heartfelt thanks to my co-advisor, Ellen Do, for the important role she played in my Ph.D. experience. Ellen advised the initial Drawing Apprentice prototype and has been a wellspring of inspiration for exploring its applications and use throughout its development. More importantly, her creative, imaginative, and joyous spirit helped me develop a new perspective on developing creative technologies that benefit the world. In addition to her academic guidance, I am sincerely grateful that Ellen continuously took the time to share valuable wisdom with me about maintaining a harmonious life during the Ph.D. process.

I would also like to thank the rest of my thesis committee: Michael Nitsche, who

believed in my work and entertained long discussions about enaction, providing guidance about how to communicate these concepts in the dissertation. Mark Riedl, who served as my initial advisor and with whom I wrote my first academic paper. Working with Mark helped me learn how to frame research and communicate findings in the field of computer science. Ashok Goel, who provided guidance about scoping my work and deriving generalizable design principles from evaluating the Drawing Apprentice system. I am also grateful to Mary Lou Maher for offering a new perspective to my work based on her distinguished career designing and evaluating creative systems.

The work presented in this dissertation has been a highly collaborative endeavor, which seems fitting for a dissertation that focuses on collaboration. First and foremost among my collaborators, I would like to thank Chih-Pin Hsiao, who has been my research partner for the better part of five years. During this time, we became close friends, worked together on practically all of our research endeavors, and co-authored numerous papers. To me, Chih-Pin is like the mountain, standing grounded and wise. Though sometimes my imaginings wandered to places far and wide, Chih-Pin would come along and gently guide me back to relevance, ensuring we bring back any valuable insights. While words cannot repay his noble efforts, I offer my humble thanks for his friendship, guidance, and work on the Drawing Apprentice system and creative sense-making theory.

A special thanks also goes to Kunwar Yashraj Singh for his tireless efforts exploring cutting edge interactive machine learning techniques in the Drawing Apprentice and Pretend Play projects. Yashs passion for cognitive computing sparked many enlightening discussions and moved my work forward tremendously. It was Yash that first pointed out the link between sense-making and the free-energy principle of the brain that eventually became a cornerstone of this dissertation.

I would also like to thank all the students that have worked on the Drawing Apprentice (undergraduate students: Brenda Lin, Lily Lau, Ricardo Macias, Raphael

Gontijo Lopes, Sanat Moningi, Justin Le, Amy Zhuang, Cora Wilson, Monet Tomioka; Masters students: Lisa Li, Ivan Sysoev, Dingtian Zhang, Mukul Sati) and Pretend Play project (undergraduate students: Margeaux Comerford, Margaret Hu, Yue Ting Lee, Shan Li, Raphael Gontijo Lopes, and M.S. students: Shweta Raj, Jordan Movish, Bradlynn Walker) in some capacity over the years. Thanks also to Caleb Southern for providing helpful feedback and camaraderie throughout the program.

Mentors from my past also deserve special thanks. In particular, I would like to thank Yanna Popova for encouraging me to continue my early work investigating perception and creativity (back in 2007 at Case Western Reserve University), which eventually evolved, with her continued input, into the creative sense-making framework presented in this thesis. I would also like to thank Per Aage Brandt, whose mentorship first introduced me to building cognitive theories, including how to root out inconsistencies and refine a theory through collaboration. Thanks to Florin Berindeanu for teaching me about the less discussed, but hugely important, mystical side of knowledge generation and truth finding. Without this dynamic trio supporting my early research, I would have never developed the passion for cognitive science and creativity that lead me to pursue my graduate studies.

Last, but not least, I would like to thank my parents, Wendy Phelps and Mark Davis, who gave me life and supported my journey to this point. I am grateful for the selfless sacrifices my mother made to ensure I was able to pursue higher education. Her life is a tribute to the idea of mind over matter, overcoming the limitations Multiple Sclerosis placed on her by learning to rewire her brain and distribute her cognition onto the environment. She taught me that anything can be accomplished through will and determination, regardless of external circumstances. I am also grateful to my father for supporting my creative inclinations from a young age, helping me build whatever I dreamed up. My dad helped me experientially understand human creativity. Spending time with him in the houses he built was like dwelling inside an extension of his mind, externalized for the benefit of others. Witnessing his creative passion and its fruits inspired me to want to help others develop their own creative and artistic passions. The tremendous support of my whole family has been invaluable, especially my sister, Jamie Davis, what has always been there for me, my Granny, Kathleen Davis, who always offered sage advice about my life, and my Aunt and Uncle, Cheryl and Mike Tully, whose kindness and good nature helped teach me what family means.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The field of computational creativity is beginning to investigate how co-creative agents might interface with the human creative process. These *computer colleagues* are a mix between creativity support tools helping users achieve creative goals and creative algorithms that generate content autonomously. Computer colleagues have enormous potential because during creative improvisational collaboration, a new form of distributed creativity arises that can lead to *emergent, dynamic, and unexpected meaning* to support creativity in new ways. However, there is a gap in the literature about cognitive accounts of the interaction dynamics of open-ended creative collaboration, e.g. the rhythm of interaction, style of turn taking, and manner in which participants are mutually making sense of a situation. An empirically grounded cognitive framework would greatly aid in the design and evaluation of co-creative systems. With this dissertation, I begin to address that gap by asking the overarching research question: *How do humans collaborate in open-ended improvisational creativity, and how can we design co-creative agents to achieve similar benefits as human collaboration?*

The thesis statement is that co-creative agents, such as collaborative drawing partners, can inspire new ideas and motivate interaction during open-ended and improvisational creative collaboration on a shared canvas. Additionally, I claim using participatory sense-making as a theoretical lens to model and quantify co-creation (e.g. interaction dynamics, emergent meaning, coupling, autonomy) can help objectively evaluate the effectiveness of interaction designs and technical approaches in co-creative systems. The methods for exploring these claims are:

- *Empirical Investigation:* Performing experiments to study the open-ended improvisational collaboration between humans in the rich domain of narrative-based pretend object play and using participatory sense-making to characterize collaboration dynamics.

- *Modeling Interaction Dynamics:* Creating a new qualitative data coding technique to perform continuous interaction analysis. Using this technique to classify different types of sense-making strategies and collaboration styles; testing the validity of these classifications through inter-rater reliability and comparison to the ground truth of the observational data.

- *Design and Evaluate a Technical Probe:* Developing a co-creative drawing agent based on the core principles of participatory sense-making, such as real-time feedback and dynamic meaning construction, to evaluate the effectiveness of different machine learning approaches, interaction designs, and interface design techniques by testing the system with users on open-ended collaborative drawing experiences with the co-creative drawing agent and wizard of oz setups.

This dissertation extends the cognitive science theory of enaction and participatory sense-making to the domain of open-ended creative collaboration to formalize this theory in computational models of creative collaboration. This knowledge primarily contributes to the fields of computational creativity, human-computer interaction, cognitive science, and creativity research. The creative sense-making framework proides a new method to rapidly and reliably quantify interaction dynamics such that they can be mathematically analyzed using continuous functions (i.e. moving averages, integrations) to understand how collaborations flow through time (versus the typical discrete, event-based qualitative coding and descriptive statistics). Finally, this dissertation produced a web-based co-creative drawing agent, the Drawing Apprentice, that learns through interaction and can serve as an experimental platform for studying different techniques of interactive machine learning, human-computer collaboration, and human-human collaboration.

CHAPTER I

INTRODUCTION

The field of computational creativity is beginning to investigate how co-creative agents might interface with the human creative process. These computer colleagues are a mix between creativity support tools helping users achieve creative goals and creative algorithms that generate content autonomously [32]. Developing computer colleagues that can collaborate and improvise in open-ended creative domains is a significant and important challenge for artificial intelligence (AI) and cognitive science research [103]. Recent advances in machine learning, such as deep learning, have resulted in extremely powerful algorithms for generating artistic content [69, 78, 155]. As a result, understanding effective ways to implement these powerful machine learning advances in co-creative systems is becoming a very active research area, especially as AI continues to gain in popularity and use.

A large portion of computational creativity research has focused on generative creativity, i.e. developing systems that autonomously generate products that people would judge as creative [21, 23, 26, 69, 137, 172, 173, 179, 180]. A growing contingent of work is beginning to develop *co-creative systems* that can collaborate with users in real time on open-ended creative tasks [9, 25, 30, 31, 89, 94, 95, 103, 105]. They have enormous potential because during creative improvisational collaboration, a new form of distributed creativity arises that can lead to emergent, dynamic, and unexpected meaning to support creativity in new ways [146]. However, designing for collaboration in open-ended creative domains presents novel and thorny challenges around how such an interaction should be carried out and evaluated [12].

There is a gap in the literature about cognitive accounts of the interaction dynamics of open-ended creative collaboration, e.g. the rhythm of interaction, style of turn taking, and manner in which participants are mutually making sense of a situation. Such a framework would greatly aid in the design and evaluation of co-creative systems. With this thesis, I begin to address that gap by asking the overarching research question: *How do humans collaborate in open-ended improvisational creativity, and how can we design co-creative agents to achieve similar benefits as human collaboration?*

1.1 Personal Motivation

My personal motivation for investigating this domain is my extensive practice-based experience in the domain of collaborative drawing over the past 15 years. During my experiences collaborating with many individuals of different experience levels, I consistently saw how turn-taking based open-ended collaboration can help both experts and novices. This realization inspired me to create a computer colleague that might have similar benefits as I was able to achieve through collaboration, such as:

Lower barrier of entry: Novices seem to feel more comfortable engaging in the artistic creative process when they are only expected to draw a few lines per turn in a collaboration. The subjective value attached to any individual line in a turn is reduced due to the highly mutable state of the artwork over the course of many turns. As a result of this value change, the experience shifts from trying to represent a particular idea (which is difficult) towards actively participating in a gradual unfoldment of evolving ideas (which is easier).

Creative inspiration: Each contribution inspires and invites the participant to respond to it, making an addition or possibly resolving some tensions that arose when an expectation was violated (thus inspiring new ideas). Often, the end result is more creative than what could have been accomplished individually due to the diversity of ideas, skills, styles, and artistic visions.

Motivation to continue: Collaboration motivates individuals to continue by shifting the value of engaging in art-making from: A) producing a good final product, to B) making interesting contributions in response to your partner through time. The dialogical element of the collaboration seems to compel the interaction to move forward, like participating in a conversation that naturally unfolds through time.

Given these experiences and an understanding of the academic landscape of computational creativity, my dissertation work focused on designing, developing, and evaluating a co-creative drawing partner, the Drawing Apprentice. Along the way, I studied human collaboration to inform the design and evaluation of the Drawing Apprentice system. However, since co-creation is such a new domain, customized research methods and tools naturally emerged during the course of evaluation. While the main contribution of this thesis is the design and evaluation of the Drawing Apprentice system, the research tools and cognitive theories presented here also have significant value to the study of co-creation more broadly.

1.2 Technical Need

The emergent nature of social interactions in open-ended creative domains presents significant challenges for traditional approaches to computational creativity since they are focused largely on enabling a system to generate a creative product or simulate the potential cognitive process by which those products are typically produced [22]. This is due partly to creativity research in general similarly emphasizing either the creative person, product, process, or environment [5] (rather than a person or people situated in an environment interacting to gradually build meaning and incrementally producing artifacts based on dynamic feedback). Designing systems to facilitate open-ended improvisational creative collaboration requires understanding the interaction dynamics of the creative and collaborative process, such as how meaning and ways of interacting emerge in the moment and serve to shape and guide creative interactions moving forward.

In the past few decades, a new theoretical paradigm has gained popularity in cognitive science called *enaction* [162]. This new cognitive paradigm posits social coordination and interaction as the basis for cognition and meaning making [41]. This thesis proposes that the enaction paradigm, and in particular its conceptual framework of participatory sense-making, can serve as a solid theoretical foundation for research into creative collaboration and co-creative systems. This new theory can help researchers understand how humans collectively make sense of creative tasks, as well as how co-creative systems and human-robot systems might effectively engage users in these types of open-ended contexts [65, 174, 175].

In the enactivist paradigm, researchers apply two main forms of interaction analysis to understand and model sense-making and participatory sense-making:

- *Designing simplified virtual environments* that provide quantitative insight into highly constrained yet open-ended social interactions. For example, the experimental setup called perceptual crossing has been recently used to quantify fine-grained interactions during particiaptory sense-making. [2, 39, 41, 42, 62].

- *Qualitative coding analysis* that provide detailed qualitative insights into the cognitive processes and mechanisms involved in more naturalistic and complex open-ended interaction [7, 28, 141].

There is a gap in research methodologies for formally investigating and modelling the interaction dynamics of naturalistic and open-ended social coordination in a quantitative way. This gap in the research is particularly important given that such a technique and the accompanying knowledge generated through its application could be leveraged to design co-creative systems as well as analyze the effectiveness of these systems.

1.3 Hypotheses

- A co-creative drawing partner that draws with users in a turn taking manner can engage users in a process of participatory sense-making (e.g. emergent meaning, coupled interaction) to offer similar benefits as human collaboration, such as inspiring new ideas and increasing creative engagement through interactive dialogue.

- The cognitive science theory of participatory sense-making can be productively mapped to the field of co-creation to provide a solid cognitive paradigm and framework to understand the unique sociocognitive elements of co-creation.

- Quantifying the interaction dynamics of co-creation using the principles of sense-making can provide more targeted data with which to evaluate the interaction design and technical approach of co-creative systems.

1.4 Thesis Statement

Co-creative agents, such as collaborative drawing partners, can inspire new ideas and motivate interaction during open-ended and improvisational creative collaboration on a shared canvas. Using participatory sense-making as a theoretical lens to model and quantify co-creation (e.g. cognitive states, interaction dynamics, emergent meaning, coupling, autonomy) can help objectively evaluate the effectiveness of interaction designs and technical approaches in co-creative systems

1.5 Research Questions

This research program asks the overarching research question: /textithow do humans collaborate and how can we design more effective co-creative agent?. It explores the interaction dynamics of open-ended creative improvisation to model and understand how co-creative agents can effectively engage users in collaboration. The research questions are as follows:

RQ1: *Can a co-creative drawing partner leverage user sketch input and feedback to facilitate participatory sense-making in a similar manner as humans, e.g. emergent meaning, coupled interaction, inspiring dialogical interaction?*

Introducing a co-creative agent as a collaborative partner fundamentally changes the nature of the creative activity due to the users pre-conceived notions about computational creativity and artificial intelligence. With this question, we specifically investigate how users perceive, conceptualize, and relate to co-creative agents in the context of open-ended collaborative improvisation. For example, do users attempt to teach the system? How do they evaluate whether the system is learning? What degree of intentionality do users attribute to the system, and how does the interface design affect this? Does the virtual character produce an affective bond with the agent, and does that influence the perceived creativity of the system?

Furthermore, there are many different types of machine learning and knowledge engineering that might be recruited for developing co-creative systems. With this question, we seek to understand the effect of different types of machine learning approaches and how they are presented to the user in the context of collaborating with a co-creative agent. In particular, we explore how real-time feedback might be recruited to help coordinate in-the-moment interactions and teach the agent knowledge on the fly.

RQ2: *Can the cognitive science theory of participatory sense-making be productively mapped to the field of co-creation to provide a solid cognitive paradigm and framework to understand the unique sociocognitive elements of co-creation?*

The enactive paradigm in cognitive science provides a novel lens through which to view collaboration and meaning production as an emergent phenomena resulting from in-the-moment interactions and dynamic perception-action feedback loops. Since this approach emphasizes the interactive and collaborative nature of meaning production, it is naturally well suited as a framework for understanding improvisational collaboration. However, the ideas of sense-making that enaction puts forward have primarily been used to describe low-level meaning production, such as negotiating movement and performing spatial and embodied reasoning tasks [133, 143], perceptual reasoning [42], sensory augmentation [63]. This theoretical weakness has been noted within the field [162], and more recent efforts have expanded the enactive paradigm into higher order cognitive processes, such as narrative reasoning [135], social coordination [54, 66, 128, 161], and social cognition in general [29, 41]. However, despite the obvious link between an interaction-based view of cognition and the creative process, sense-making and the enactive perspective have not been systematically applied to understanding the creative process directly, or open-ended improvisational creative collaboration in particular.

RQ3: *Does continuous data quantifying the interaction dynamics of co-creation help evaluate the technical approaches and interaction design of co-creative systems?*

With this work, we hope to lay the foundation for how the enactive perspective may be recruited for understanding the creative process in general, as well as formal methods for studying collaboration and social coordination in open-ended improvisational contexts through the lens of enaction. With this research question, we seek to identify whether a more formalized research method and cognitive modeling technique can be developed to understand open-ended creative collaboration. We propose such a technique by recruiting complimentary cognitive theories to support this formalization. Further, we ask whether this technique can yield a quantitative method of identifying and classifying different patterns of interaction dynamics during sense-making and participatory sense-making throughout collaboration.

To address this question, we apply our new analysis and modeling technique to multiple domains of creative collaboration (i.e. pretend play and collaborative drawing) and evaluate whether this approach provides a productive addition to the overall sense-making framework.

1.6 Methods and Evaluation

This thesis uses a mixed-method approach to address and explore the research questions outlined above. There are three main activities comprising the methods: 1) developing a co-creative agent to serve as technical probes to explore the implementation of these insights and evaluate the effectiveness of using the proposed conceptual framework for designing and evaluating human-computer collaboration in open-ended creative domains, 2) analyzing human collaboration to develop a systematic coding approach that combines the benefits of qualitative analysis with the computational power of quantitative analysis, and 3) applying this technique to classify and quantify different types of sense-making in open-ended creative contexts.

- *Design and Evaluate a Technical Probe:* Developing a co-creative drawing agent based on some of the core principles of enaction, such as real-time feedback and dynamic meaning construction, to evaluate the effectiveness of different machine learning approaches, interaction designs, and interface design techniques by testing the system with users on open-ended collaborative drawing experiences with the co-creative drawing agent and Wizard of Oz setups.

- *Empirical Investigation:* Performing experiments to study the open-ended domain of pretend play and conducting a qualitative coding analysis beginning with Grounded Theory to analyze the interaction dynamics and collaboration techniques used during play sessions.

- *Modeling Interaction Dynamics:* Creating a new qualitative data coding technique to perform continuous interaction analysis. Using this technique to classify different types of sense-making and collaboration styles and strategies and testing their validity through inter-rater reliability and comparison to the ground truth of the observational data.

We perform a series of evaluations to investigate the Drawing Apprentice co-creative drawing partner prototype. The Drawing Apprentice prototype is designed as a technical probe meant to begin exploring how a co-creative agent can interface with humans on an open-ended creative task, such as collaborative drawing. The system evaluation method includes both formative studies evaluating design decisions as well as summative studies exploring the effect that collaborating with the system has on the creative process of users.I rely on the conceptual framework of participatory sense-making to help characterize dimensions of collaborative interaction that centers on three concepts: *coupled interactions, emergent meaning, and creative engagement.* Coupled interaction relates to the structural correspondences between the actions of the creative agents involved in the collaboration. Actions that are coupled are mutually influential, meaning contributions from each party are clearly influential in what follows. Emergent meaning refers to seeing new meaning structures as a result of the interaction, i.e. users elaborate or transform their ideas in response to interacting with the agent. Finally, creative engagement, in the current context refers to the motivation and inspiration that users derive from the dialogical nature of interacting with a collaborative partner. The call-and-response interaction between the user and system provides opportunities for action that can propel the artwork forward by helping users think of what to do next.

In the summative user study, I compare the creative behavior and experience between in two improvisational drawing conditions described below:

- **Wizard of Oz Collaboration:** A human user collaborates with the user in a 'Wizard of Oz' style setup where the user is not aware they are collaborating with a human.

- **Drawing Apprentice Collaboration:** The user collaborates with the Drawing Apprentice system.

After each experimental condition, users engage in a retrospective protocol during which they will watch their drawing session and describe their experience interacting with the system. Additionally, users will be given a survey and answer a few questions in a semi-structured interview that focuses on their qualitative experience interacting with the system. Finally, users will be engaged in a debrief protocol during which they will be asked to compare and rank their experiences with the different systems.

To understand open-ended improvisational collaboration, we empirically investigate improvisation in multiple open-ended domains including pretend play and collaborative drawing. Pretend play was selected as a domain of inquiry due to the similarities it had with the creative process of abstract art, i.e. there is no definite goal at the onset, meanings grow and transform through time, and there are no set rules for what it means to play outside the play context. Additionally, the narrative that emerges through play is an ephemeral creative product that emphasizes in-the-moment interactions rather than long periods of reflection between actions.

A study was performed analyzing the pretend play behavior of 32 adult dyads during two five minute play sessions. Before each play session, players were provided a prompt, such as 'drag racing' or 'monsters attack' that they were told to incorporate throughout the session. The analysis investigated how meaning was formed in the play session and what types of activities influenced the relative success and creative engagement of players during their session. From this study, a cognitive model of improvisational collaboration was developed. Additionally, we developed the sense-making curve technique to further quantify interaction dynamics in open-ended improvisational collaboration. Multiple coders produced curves for this dataset (as well as the Drawing Apprentice studies described below) to ensure inter-rater reliability. Different types of participatory sense-making and collaboration were identified and the approach was validated by cross-domain application and measuring inter-rater reliability within each domain.

1.7 Contributions

This thesis work produced a co-creative drawing agent that learns through interaction and can serve as an experimental platform for studying different techniques of interactive machine learning, human-computer collaboration, and human-human collaboration. In addition to its academic utility, the drawing agent has pragmatic applications in the domain of education by creatively inspiring children's artistic, creative, and collaborative abilities as well as providing a means to quantify their coordination and collaboration practices.

The theoretical component of this thesis extends the cognitive science theory of enaction to the domain of open-ended creative collaboration and an accompanying elaboration of the theoretical framework of sense-making that helps to formalize this theory with respect to computationally modeling creative collaboration. This knowledge primarily contributes to the fields of cognitive science, creativity research, and human-computer interaction. It contributes to cognitive science by demonstrating how enaction and participatory sense-making can be productively applied to describe open-ended creative collaboration. It contributes to creativity research by providing a novel perspective on creativity and collaboration, and finally, it contributes to human-computer interaction by demonstrating how the enactive paradigm can serve as a useful cognitive foundation upon which to base design decisions related to open-ended collaboration and interaction, such as turn-taking.

The sense-making curve technique itself provides a new research method to quantify interaction dynamics such that they can be mathematically analyzed using continuous (rather than the typically discrete qualitative coding data) to understand how collaborations flow through time. The sense-making curve is a new technique for performing a hybrid qualitative/quantitative analysis of observational data and technical tools to perform such analysis. This new knowledge can be applied to develop theoretically and empirically rooted models of creative collaboration in co-creative systems, including human-human collaboration and human-computer collaboration.

1.8 Thesis Overview

Chapter 2 explores related work including theories of creativity, background about creativity support tools, computational creativity and co-creative systems. Chapter 3 introduces the idea of enaction and participatory sense-making and proposes a new view of creative sense-making that is built upon throughout the thesis. Chapter 4 describes the Drawing Apprentice system, which is the technical probe we built to explore different interaction design and machine learning techniques that help facilitate participatory sense-making in open-ended collaboration. Chapter 5 describes the user study evaluation of the Drawing Apprentice system. Chapter 6 describes the empirical investigation of creative collaboration between humans in the domain of pretend play to further operationalize the theory of participatory sense-making in creative improvisation. This chapter puts forth a conceptual characterization of pretend play using the principles of enaction, and it outlines the need for a new coding method for quantifying interaction dynamics in open ended improvisational collaboration. Chapter 7 describes the proposed sense-making curve coding technique and theoretical justification. The mathematical procedures and analysis methods are described in detail and applied to the pretend play dataset to demonstrate its utility for classifying trends in the interaction dynamics of open-ended collaboration. A sense-making curve analysis is conducted on the Drawing Apprentice user study data and the pretend play human collaboration data and analyzed to generate design principles for co-creative systems. Finally, Chapter 8 describes future work further validating the sense-making curve technique through application in EEG, as well as the technical vision for the Drawing Apprentice system.

CHAPTER II

RELATED WORK

2.1 Summary

The chapter will first briefly contextualize this research in the broader field of creativity research. After developing a basic definition of the type of creativity involved in this research, we will turn to creativity support tools, with an emphasis on drawing CSTs. Finally, we will explore a number of computational creativity projects that involve that computer as a colleague in the creative process. Although these projects are from other domains, their interaction design and issues can provide insights about the collaborative element of this research.

2.2 Creativity Research

Early creativity research focused on understanding the creative genius of influential individuals, such as Galileo or Einstein [111, 119]. These researchers identified personality traits and environmental factors that contributed to creative genius [79,168]. As research continued, however, the cognitive mechanisms underlying creative cognition indicated that creativity was not a special gift endowed only to a few, but rather it utilized everyday cognitive mechanisms in a special way [156]. In fact, a taxonomy of related concepts emerged to help guide creativity research, the creative person, process, product, and press or environment [111]. With this new perspective, researchers posit a continuum hypothesis for creativity, explaining that creativity is a a continuum ranging from everyday or little-c creativity to more advanced historically creative discoveries (big-C Creativity), such as Einstein's Relativity [118, 136] .

The type of creativity exhibited in abstract collaborative art is closer to the everyday creativity end of the continuum because artists make small improvisational decisions rather than elaborate planning activities more commonly associated with traditional representational art [104]. Because of their fluidity, decisions in abstract art continually redefine the problem space, and the creative process resembles problem-solution co-evolution observed in design creativity [48, 83, 148]. Instead of a linear path toward a solution, problem-solution co-evolution refers to a process where redefining and refining the solution restructures the design problem, and vice-versa. There are no set rules for an abstract artwork. Rules can emerge, such as patterns or themes that exhibit their own internal logic and consistency [31]. However, the rules are loosely defined and constantly evolving, similar to flexible rules and concepts in creative pretend play where children fluidly assign new meaning to elements in the environment to suit their current pretend task (e.g. cardboard lightsabers) [189]. In these fluid improvisations, the constraints for interpreting sensory data are loosened, and conceptual shifts occur, whereby changes in frame of reference cause concepts to merge and blend with those in other domains [102, 120].

In collaborative abstract art, creativity is situated, meaning decisions occur in real-time without fully planning the actions ahead of time [3, 71, 163]. Creative decisions are negotiated based on the real-time feedback and interaction with the environment (including contributions from other collaborators) [92]. Given the importance of situated creativity for our system, research will investigate the creative experience of the artist. Creative experience can only be subjectively defined, and it simply represents the embodied experience of being creative, i.e. the phenomenological experience of engaging in creativity [114]. To understand how computational collaboration affects the creative experience, we will qualitatively study the 'how' and 'what' of the Drawing Apprentice system, namely the manner in which it makes creative contributions (i.e. interaction design), and what type of creative contributions are effective, respectively.

2.3 Computational Creativity

Computational Creativity is a branch of Artificial Intelligence that focuses on creating systems that have open-ended goals and creative ways of accomplishing those goals. The field of computational creativity is dedicated to producing machines that exhibit creative behavior or perform tasks in a creative way [16, 24, 107, 130]. These computational systems exhibit creativity by producing creative artifacts independent from the programmers that created them.

Harold Cohen's painting program named Aaron is a prime example of a computationally creative system [113]. The system architecture for Aaron is important to this discussion because it relates to situated creativity. Aaron features multiple independent modules that all have independent parameters used to make creative decisions and develop their own output. It is similar to Brooks' embodied robots [13] in which intelligence is distributed among a network of sensor-actuator pairings rather than a centralized planning system. In Brooks' robots, intelligent seeming behavior emerged from the interaction of relatively un-intelligent sense-act cycles [13]. Aaron begins with a random injection, and then each module perceives the initial starting data and provides its interpretation and computational reaction to that data point. These computational reactions are artistic contributions such as shapes, figures, colors, and textures. Each of these modules makes contributions and feed into each other in a way that produces a cohesive painting. The painting is an emergent result of independent processes.

Aaron's artwork is considered non-deterministic because Cohen did not program the computer to create the exact artwork [113]. Small changes in the process can create a large difference in the end because there are many independent modules that interpret data. Each of these modules may have conflicting interpretations or different reactions to the line.

Figure 1: A painting from Harold Cohen's Aaron

2.3.1 Models of Computational Creativity

It is helpful to construct computational models of creativity to build systems that exhibit creative behavior. Boden (2003) analyzed computational creativity and generated some useful terms and categories that have helped structure the field [11]. Perhaps most notable is her psychological supposition that we can consider personal creativity as distinct from historically creative acts [11]. Boden refers to personally creative acts as psychologically creative, or p-creative. Historically creative (h-creative) acts are socially recognized and valued contributions, such as Einstein's theory of Relativity. Boden also distinguishes between different two types of creativity: combinatorial creativity and transformational creativity [11]. Combinatorial creativity works by combining existing components of a domain in a novel manner. Transformational creativity works by changing the rules or frame that constitute of a problem to devise a novel solution.

2.3.1.1 Combinatorial Creativity

When considering combination, the ideas of a search space and problem space becomes important. A problem space represents the entire collection of potential elements that are relevant to a particular situation. For example, if we are trying to invent a new vaccine or medicine, the problem space would include all the existing medicines as well as their individual components, including the chemicals that constitute them. Humans are not particularly good and performing an exhaustive search of a problem space because of the combinatorial explosion that occurs when the number of elements in a search space reaches more than a few items. Computers, on the other hand, can perform high speed calculations that can compute thousands or millions of combinations and analyze each of the outcomes. In fact, computer programs have helped discover new chemical composition using this method of cycling through all possible combinations to try to find the optimal solution (i.e. the brute force approach) [11]. One could describe Aaron as exhibiting combinatorial creativity because the system works within a set of rules and constraints. Within each of those rules, randomness is interjected to introduce novelty, but the rules themselves do not change. Approaches have also been developed to computationally evaluate this type of novelty [77, 109].

2.3.1.2 Transformational Creativity

The other type of creativity Boden describes is called transformational creativity, which fundamentally changes the problem space. This type of creativity transforms the space of possibilities by re-interpeting the problem or shifting its application in some manner. Computers have a more difficult time with this type of creativity because it often entails analogical reasoning. In analogical reasoning, the structure of a source domain is used to reason about the structure of a target domain [53,90,182].

Koestler described this type of transformational creativity as the intersection of two thought matrices, or sets of domain knowledge [99]. The difficult part of this process is identifying two domains that can be spliced together to yield an insight into a problem. Problems that require this type of thinking are referred to as insight problems. To solve an insight problem, one must shift the frame of reference for thinking and apply an ulterior set of logic and assumptions [151].

In collaborative abstract art, transformational creativity occurs when the shapes and figures of the artwork are re-interpeted and shift meaning throughout the creative process. Since there is no explicit goal at the onset, the intention of the artist remains fluid enough that conceptual shifts may not be as difficult as in problem solving. Insight problems are difficult because humans fall prey to what is referred to as fixation [186], where we concentrate on one particular type of interpretation or frame of reference for a problem, which makes it difficult to interpret that problem in any other way. I predict that injecting computational contributions to an artwork might help the process remain fluid and open to incorporating new ideas, thus yielding a more flexible and creative product.

2.4 Creativity Support Tools

Once it was established that creativity can be trained, facilitated, and measured, researchers began to develop techniques to support creativity [27, 79, 156] . Initially, these techniques were procedural activities one could engage in to stimulate creativity, such as brainstorming and lateral thinking exercises [37,138]. Researchers also began developing a class of technology referred to as Creativity Support Tools (CSTs) [17, 84,152,154] CSTs are designed to help users explore a creative domain, record decision histories, and scaffold skills to allow and encourage users to learn expertise [15,153].

Shneiderman (2007) distinguishes creativity support tools (CSTs) from productivity support tools through three criteria: clarity of task domain and requirements, clarity of success measures, and nature of the user base [153]. Productivity support tools are designed around a clear task with known requirements, have well-defined success metrics, and are characterized by a known and relatively well-understood set of users. In contrast, CSTs often work in ill-defined domains with yet-unknown or unknowable requirements, have vague success measures, and have an unknown user base or one that behaves in unorthodox manners. For example, consider support tools for the well-defined goals of product supply scheduling, which include many clearly defined variables like cost metrics for shipping efficiency. Contrast this with a drawing support tool. Here, task requirements consist of a plethora of elements that define a drawing, success is measured by user acceptance of their final product.

Creativity support tools can take many forms. Nakakoji (2006) organizes the range of creativity support tools with three metaphors: running shoes, dumbbells, and skis [116]. Running shoes improve the abilities of users to execute a creative task they are already capable of; they improve the results users get from a given set of abilities. Dumbbells support users learning about a domain to become capable without the tool itself; they build users' knowledge and abilities. Skis provide users with new experiences of creative tasks that were previously impossible; they enable new forms of execution. A contemporary text editor that highlights grammar mistakes is a running shoe; explaining why those wordings are ungrammatical makes the tool a dumbbell. Collaborative drawing tools are a type of ski because they enable a whole new class of creative expression where the user collaborates with a computer.

Nakakoji believes CSTs that introduce new creative experiences to novices will gain popularity because of the positive impact novel creative experiences can have on creative output [116]. However, we do not currently have an adequate understanding of these tools [116].

2.4.1 Design Creativity Support Tools

Many tools and techniques have been developed to support creativity during the design process, such as supporting the expression and exploration of many design ideas [181] and helping to inspire ideation by providing relevant information. For example, the DANE system provides biologically inspired examples that are analogically related to the designer's current problem or solution domain [171]. This type of tool has the capacity to help designer's discover novel and surprising solutions based on analogically related concepts, which has proven useful in educational contexts [96, 184].

Extensive research has also been done on exploring methods to integrate machine learning and artificial intelligence into creativity support tools to assist designers and artists achieve their creative goals more effectively. For example, Do and Gross' Electronic Cocktail Napkin project [45] tries to understand symbolic and spatial relations of the user's early design sketches to help the system understand the content of the user's drawing, such as the type of object, its components, and their spatial relationship. Similarly, the Right-Tool-Right-Time analyzed the user's design sketch to dynamically offer different types of creativity support based on what type of task the user is currently engaged in [45].

The SolidSketch system extended Do's initial concept of providing contextually grounded support in creative design tasks from a 2D environment to a 3D environment [91]. This sketch-based tool allows users to rapidly sketch out 3D object using context dependent gestural commands to manipulate the content of the 3D model. Mechanix is another intelligent design creativity support tool targeted for education that analyzes the user's sketch and creates realistic engineering models and simulations based on features of the drawn sketch [117].

2.4.2 Drawing Support Tools

There are a variety of CSTs focused on supporting artistic creativity. Many of these CSTs try to increase the technical ability of artists, such as ShadowSketch, iCanDraw, and Projector Guided Painting. All these projects try to help the user draw more skillfully. Other approaches try to increase creative expression, such as Everyone Loves Sketch that enables users to draw in 3D. I am not aware of any drawing-based CSTs that collaborate with the user as an equal in the creative process, like the proposed Drawing Apprentice system.

2.4.2.1 ShadowDraw

ShadowDraw is a tool that searches a database of drawings similar to the current one and displays a composite image of similar drawings [101]. The similar images found in the visual search are layered on top of each other and placed on a digital canvas underneath the user's current drawing [101]. For example, if a user began to draw two circles for a bicycle, the system will find all images with two circles and superimpose those images with details of the bike frame, spokes, etc. Those images will appear as a 'shadow' behind the user's drawing, with the hope that some of the shadow lines can aid the user in drawing a more detailed or accurate image. This system is meant to support sketch artists that know how to draw some lines of an object, but not necessarily the entire thing. It also offers additional details that the user may not have been aware of for the object s/he chose to draw. ShadowDraw is one approach to try to increase the technical ability of sketch artists.

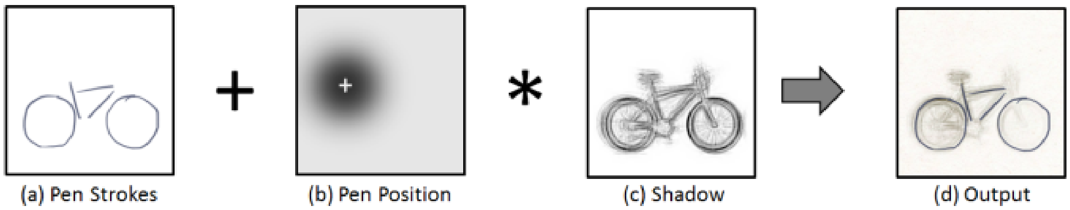


Figure 2: ShadowDraw interaction [101]

2.4.2.2 iCanDraw

iCanDraw is a drawing critique and feedback tool that helps users refine and hone their portrait drawing skills [44]. The program has an expertly drawn sketch hardcoded in its memory. As the user attempts to draw the same portrait, the system informs the users which lines are accurate and which lines are not by comparing the ground truth image to the user's drawn lines. Inaccurate lines are highlighted, which allows the user a chance to rectify the lines and make their drawing more similar to the expertly drawn image. iCanDraw strives to enable novices to create high quality portraits. The concept behind this approach could also be extended to other representational objects, such as houses or animals.

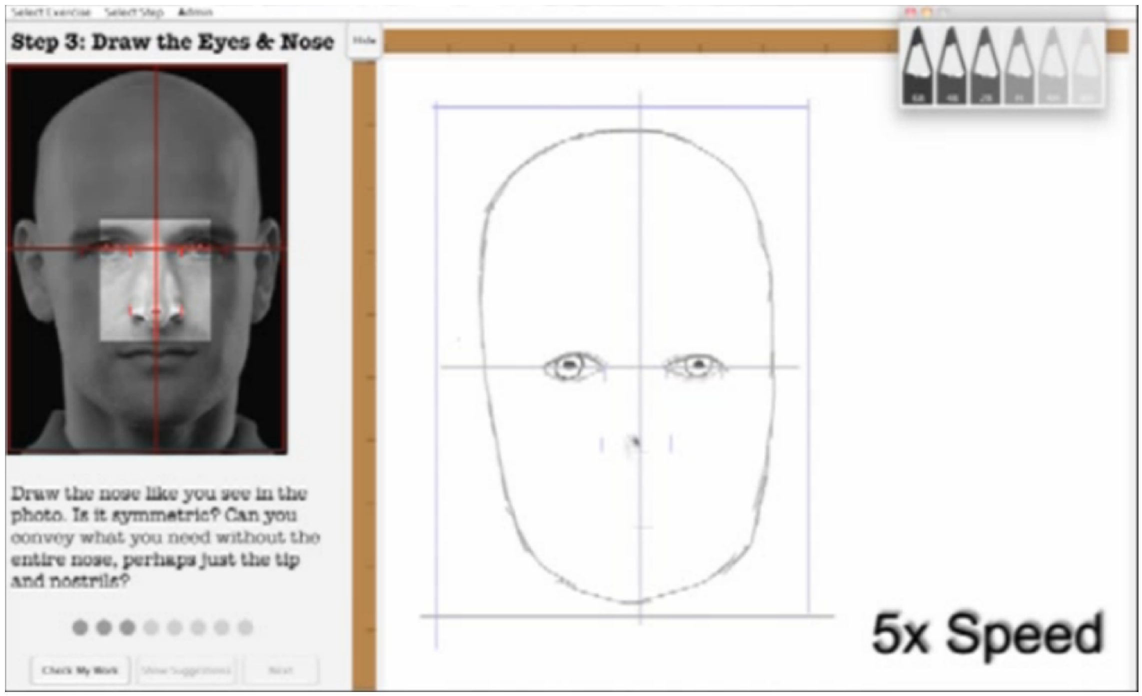


Figure 3: iCanDraw? [44]

2.4.2.3 Everybody Loves Sketch

Everybody Loves Sketch is a program that facilitates drawing in 3 dimensions [4]. Typically canvases exist on a 2-dimensional plane. Drawing techniques such as perspective and foreshortening can help create the illusion of 3D space, but the drawing ultimately always happens in 2D space. Everybody Loves Sketch allows the user to rotate the drawing plane to work on slices at different planes in the 3D space. There are also operations that help facilitate 3D drawing, such as mirroring the user's strokes to help build a solid object. Everybody Loves Sketch makes it easier to make 3D drawings, which is important given the rise of 3D printing and the need to create 3D models in a straightforward manner. The system helps facilitate a technical drawing ability, but it doesn't actually make creative contributions of its own.

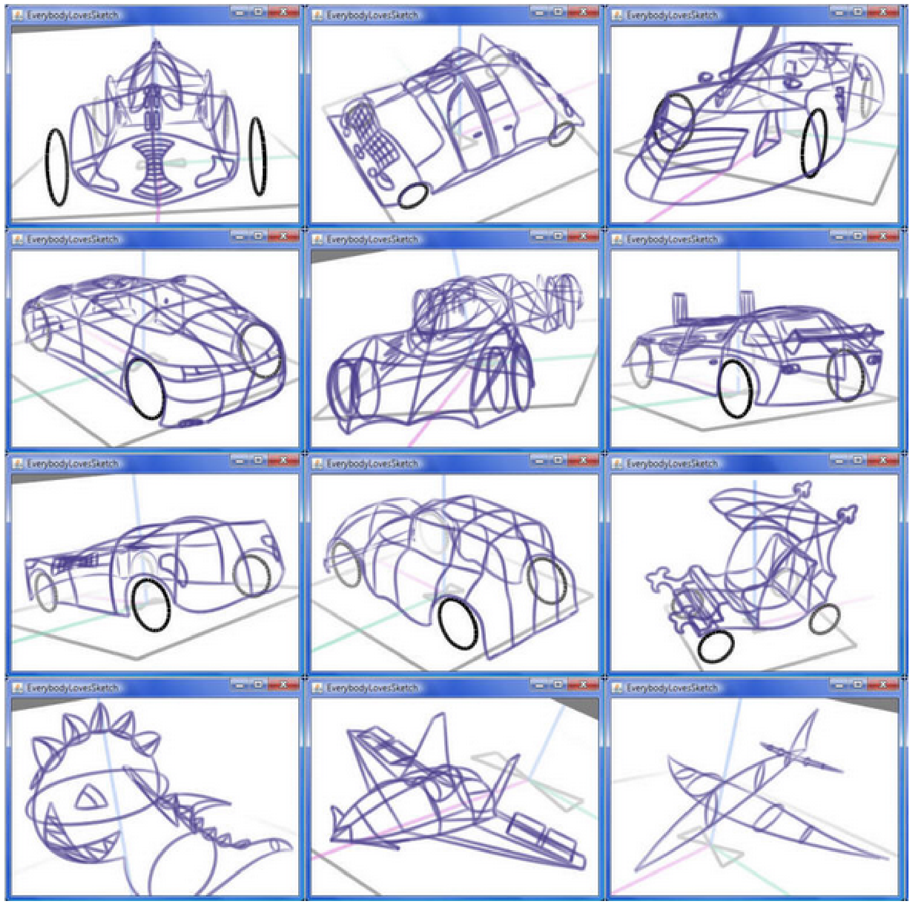


Figure 4: Everybody Loves Sketch high school student examples [4]; facilitates 3D sketching.

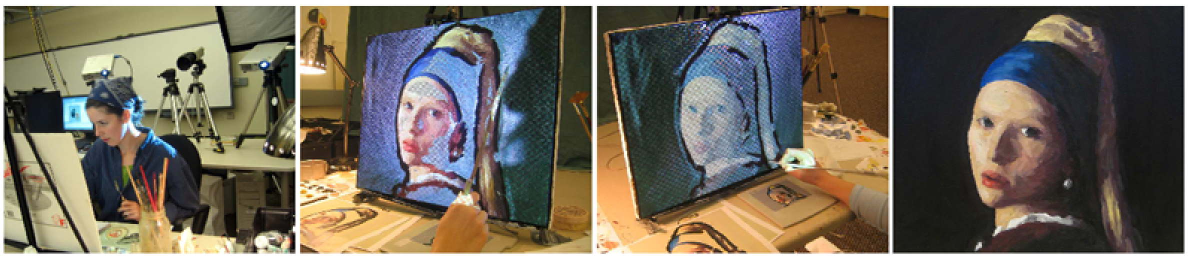
2.4.2.4 SwarmSketch

SwarmSketch is a project started by Peter Edmunds that explores the idea of collective drawing [50]. Swarmsketch.com provides a shared canvas to which anyone can contribute. The service selects a prompt that users are supposed to follow. For example, at the time of this writing, the prompt is "Edvard Much." Each user is limited to a certain number of lines to help capture the essence of the prompt. Each drawing continues for a week or until it reaches 1000 lines, whichever comes first. SwarmSketch explores the idea of crowd sourcing for creative products. Crowd sourcing is a new computational technique that asks the 'crowd' on the internet to each contribute a small amount to a project. In this case, each member of the crowd contributes a few line to a drawing. In a sense, SwarmSketch explores the idea of artistic collaboration, but that collaboration is all done asynchronously.

Figure 5: SwarmSketch [50]

2.4.2.5 Projector Guided Painting

The Projector Guided Painting project attempts to teach novices how paint like the Old Masters [56]. This system has two components, a projection system that projects a type of paint-by-numbers instruction on a real canvas, and a color mixing system that helps the user mix the proper colors for the projected instructions. There are multiple layers of projections that each represent one step in the painting process. The first projection, for example, might contain large dark shapes that help develop the base layer for the painting. The next projection might depict lighter colors and more detailed shapes that start to actually shape the content of the painting. In this way, the user is stepped through the process of creating a real oil painting. This project is interesting because it retains the medium of oil painting, whereas many other artistic creativity support tools force their users to work digitally.



(a) Painting studio setup (b) Painting a layer in preview mode (c) Painting the black regions of a layer in color selection mode (d) Final painting created using our system

Figure 6: Projector Guided Painting [56]

2.5 Human Computer Creativity

The CSTs we have looked at so far all support the artistic creative process in some way, but none of those systems contribute as an equal to the creative process of the user. Next, we will examine some projects from different creative domains that all enable the computer to be a colleague in the creative process in some manner. Although none of them deal directly with drawing (because there are no known examples in that domain), they can provide important insight into the interaction designs that facilitate human-computer creativity.

The first topic to consider is co-creativity. Co-creativity occurs when multiple parties contribute to the creative process in a blended manner [110]. In alternative creative situations, tasks can be accomplished through a distribution of labor, but the result only represents the sum of each individual contribution [110]. Co-creativity goes beyond this division of labor model and allows all individuals to contribute collaboratively and synthetically. In this situation, ideas can be fused, combined, merged, and added onto in ways that stem from the unique mix of personalities and motivations of the team members [110]. It can yield more creative solutions than if each party completed an isolated task and then added them together, i.e. the sum is greater than the parts [110].

Human-computer creativity is an interaction paradigm present in a new type of creativity support tool that introduces the computer colleague as an equal in a collaborative creative process. Depending on the implementation details, the computer can potentially collaborate with the user in a variety of ways. The crucial point here is that the computer does not follow a pre-defined script to guide the interaction. The program is adapting to the input of the user and generating responses to that input based on computationally creative algorithms. A few examples will help elucidate the variety of ways that humans and computers can co-create.

2.5.1 Interactive Improvisational Robotics

Gil Weinberg's research in interactive improvisational percussion robots provides lessons for building an improvisational drawing system [89]. Weinberg's robot, Shimon, listens to a performer and mimics or adds to the performance of the human collaborator. The system analyzes the music of performers and generates melodic improvisations that are in sync with human collaborators. In practice, the human and robot develop a call and response interaction where each party modifies and builds on the previous contribution. Improvisational creativity more closely resembles a dialogue where each party can make contributions that feed into an interactive creative process [145]. Improvisational creativity is distinguished from other types of creativity in the sense that the product is usually ephemeral; the process is the product. Jazz music has received much attention as a prototype of improvisational collaboration. creativity [145]. One source of creativity in improvisation comes from methods of responding to the contributions of others. Jazz musicians work together to form musical themes and patterns through a process of negotiation and experimentation [145]. Shimon adapts to human music, and it also has preprogrammed crescendo generating algorithms to add some dynamism to the interaction. See Figure 7 for a photograph of Shimon.

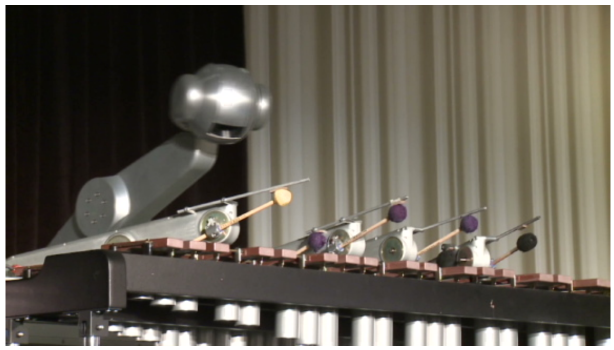


Figure 7: Gil Weinberg's jazz improv robot Shimon

Weinberg's work helps to inform the concept of creative trajectories in this thesis. A creative trajectory is a timeline dictating the sequence of creative strategies that will be executed during a collaboration. The Drawing Apprentice system has different types of creative trajectories and ways of controlling how the system decides creative strategies (i.e. manual control, hardcoded, and adaptive).

The gestural component of Weinberg's research shows the importance of a perceived relationship between the human and the computer colleague. Shimon moves with the beat in a stylistic type of dance. Currently, there is no equivalent notion of a beat for drawing. It is conceivable that the rate with which lines are drawn could represent a type of tempo. More pragmatically, and for earlier stages of development, Apprentice will achieve gestural expression in the way it draws lines. One method for achieving this is showing lines drawn like a human would draw them instead of a line that instantly appears. The system should show the gestural components of a line. As part of formative design, experiments will be conducted that modulate the speed with which lines are drawn to increase the expressivity of the system.

2.5.2 GenJam: Interactive Improvisational Jazz System

GenJam is an improvisational jazz program that Al Biles has been developing the last 20 years [9, 10]. GenJam uses a form of genetic algorithms to generate jazz improvisations. The system has acoustic sensors that recognize musical input. A number of jazz improvisation schemes are stored in the program. An accompaniment is selected by the computer and played on a MIDI style system. Biles has performed with GenJam in jazz clubs successfully for several years. He advocates evaluating such systems based on the degree to which the system is useful and meaningful to the artist.

The system has acoustic sensors that recognize musical input. A number of jazz improvisation schemes are stored in the program. An accompaniment is selected by the computer and played on a MIDI style system. Biles has performed with GenJam in jazz clubs successfully for a number of years. He advocates an evaluation scheme that includes using the system in a meaningful way. The measure of success depends on the degree to which the system is useful and meaningful to the artist. The GenJam system helps Biles perform and contributes as an equal in creative jazz improvisation. The system can be called a success according to the value it added to the creative activity.

2.5.3 Dance Improvisation

Viewpoints AI (VAI) is a co-creative dance partner that improvises with users in real time as they dance in front of a virtual character projected on a large display screen [94]. Using a Kinect, VAI analyzes the dance gestures of users and selects a complimentary dance move for the virtual character to perform. The system was initially trained by analyzing professional dancers to seed its knowledge base with expert dance moves [95]. Its knowledge base grows as it observes novel moves performed by users.

VAI has a similar goal as the Drawing Apprentice to coordinate with users during open-ended creative improvisation. However, drawing results in a creative product that remains visually present and grows over time versus the ephemeral activity of dancing.

2.5.4 Computational Theatre Improvisation

Another example of human-computer co-creativity is Magerko et al. digital improv project [105]. In this project, users engage in multiple theatrical improvisation games focused on narrative construction with AI improvisers. In the Digital Improv project, a computationally creative system attempts to recognize actors actions, interpret them to be relevant to kinds of characters, and respond to these actions in a virtual world. The computer develops a shared meaning with the actors and generates actions according to features of the negotiated narrative context [85–88].

Other actors try to guess what that action is and play along. In the digital improv project, a computationally creative system attempts to recognize actors' actions, interpret them, and respond to these actions in a virtual world. The computer develops a shared meaning with the actors and generates actions according to features of the negotiated narrative context.

2.5.5 Neural Style Blending

Recent innovations in machine learning have introduced a new class of computationally creative algorithms that have been leveraged in the creative process of the user in novel ways. For example, convolutional neural nets have been used to learn the artistic style of paintings [70]. Once the style is learned, it can be applied to a target image. The algorithm attempts to minimize the loss between the learned source image and the target image, thereby applying the textures from the learned source image to the target image. After several epochs, the algorithm gradually optimizes the style transfer, creating a new image that blends the style from the source with the content from the target.

This ML approach has become highly influential and spurred many creative applications, such as exploring how to involve users in the loop throughout this style blending process. For example, [18] explored how users could rough in a colored shapes representing a scene, which can then be stylized using a learned style. These projects are beginning to explore how computationally creative algorithms can be integrated into the creative process of users by contributing content to a shared creative product.

2.5.6 Mixed-Initiative Systems

In the world of game design, researchers have been working on what they call mixed-initiative systems that assist humans in authoring games, such as level design and narrative generation [183]. In mixed initiative systems, the human and computer are both contributing, but the contributions from the human and the computer can be quite different. In some instances, the human is generating content, and the system modifies and transforms that content.

Yannakakis et al. (2015) developed a framework for evaluating mixed-initiative co-creative systems in the context of game level design and used it to evaluate a mixed initiative game level design system called the Sentient Sketchbook [183]. These authors argue that co-creative systems require a different type of evaluation than more traditional CSTs given that they are meant to support and facilitate the actual creative process rather than solely helping to produce a higher quality creative process. Typical CST evaluation would include judging the novelty and value of the creative product.

With mixed iniative co-creativity, however, Yannakakis et al. argue there is a pressing need to evaluate the creative process toward the final outcome in addition to the outcome. They propose two metrics to evaluate the system's impact on the creative process: (1) degree of use (i.e. how often the user employed the system), and (2) quality of use (i.e. how influential and impactful the system's contributions were). With these two metrics, the authors quantify the frequency of interaction with the system and how the content of that interaction was used in the creative process.

Mixed-initiative co-creative systems bear a strong resemblence to our Drawing Apprentice prototype. However, the type of system described by Yannakakis does not grant the system the role of a partner with equal responsibility for contributing to the creative product. The system is used to help simulate, explore, and ideate around different design possibilities, but it does not directly contribute to the final creative product. Additionally, game design is not open-ended as their are hard constraints with respect to how end users of the game need to interact with the level itself.

When considering open-ended creative improvisation and collaboration with a co-creative partner, the interaction that emerges between the system and user becomes critically important. For example, the degree that Yannakakis et al. employed to evaluate their system is not wholly relevant since our system performs creative drawing actions of its own accord, similar to a human collaborator, rather than executing a task upon command. The other metric employed by Yannakakis et al., namely quality of use, is encapsulated in our evaluation in terms of quantifying whether each individual contribution from the system was effective, i.e. whether the user built upon it or integrated into their existing artistic ideas in some manner.

2.6 Conclusions

The Drawing Apprentice prototype is the first attempt to create a computer colleague in the field of visual art. There are several examples of computer colleagues from other domains, such as jazz and theatre improvisation. Additionally, mixed-initiative systems have explored some aspects of how humans and computers might both contribute to the user's creative process.

However, open-ended improvisational collaboration introduces a number of variables that are not fully explored by previous approaches, most notably the interaction dynamics that can emerge between a creative user and a creative system that both have the capacity to contribute to the creative product as equals. To evaluate such systems, we draw upon the framework of participatory sense-making that describes how multiple agents work together to define dynamic meaning structures that serve to guide their collaboration going forward. In particular, we have identified three critical areas of evaluation to determine how well a co-creative agent engages users in participatory sense-making: coupled interaction, emergent meaning, and creative engagement.

CHAPTER III

CREATIVITY AS A SENSE-MAKING PROCESS: AN ENACTIVE VIEW OF CREATIVE COGNITION

3.1 Summary

This chapter proposes a theoretical framework and modeling scheme to help understand creativity and collaboration in order to support the design and evaluation of co-creative systems. We argue that the traditional cognitive science theories currently utilized in computational creativity can be expanded to emphasize the embodied, situated, and distributed nature of the human creative process. We outline a path forward for computational creativity that involves reframing our understanding of creativity using a new theory of cognitive science called enaction that examines cognition through the lens of autonomous agents interacting in a world through a process of sense-making. A new enactive model of creativity is described along with theoretical implications for the design and functionality of computer colleagues. Introduction

3.2 Introduction

The modern landscape of computing has rapidly evolved with breakthroughs in new input modalities and interaction designs, but the fundamental model of humans giving commands to computers is still largely dominant. A small but growing number of projects in the computational creativity field are beginning to study and build creative computers that are able to collaborate with human users as partners by simulating, to various degrees, the collaboration that naturally occurs between humans in creative domains [9,33,89,103,189]. If this endeavor proves successful, the implications for HCI and the field of computing in general could be significant. Creative computers could understand and work alongside humans in a new hybrid form of human-computer co-creativity that could inspire, motivate, and perhaps even teach creativity to human users through collaboration.

To reach this optimistic future, the field of computational creativity needs a conceptual framework and model of creativity that can account for the collaborative and improvisational nature of human creativity. Traditional information processing views of cognitive science describe cognition (and resultingly creativity) as an abstracted manipulation of symbols occurring largely in the brain [123]. The new cognitive science theory of enaction claims previous theories do not properly consider the role that interaction plays in cognition. The strong claim of enactivists is that cognition (and resultingly creativity) emerges through a real-time and improvised interaction with the environment and other agents in that environment [162, 170]. The interaction with the environment is critical in order to fully undestand cognition. While traditional information processing cognitive theories could work to incorporate the perception-action feedback loop emphasized by enaction to model continuous improvised interaction, the enaction theory begins with the assumption that all cognition is based on this principle of improvised interactions guided by feedback from the environment. Since the theory of enaction emphasizes interaction as crucial to cognition, there are concepts and vocabulary that help provide nuanced descriptions of how individuals build meaning together (e.g. interaction dynamics, emergent meaning, autonomy, etc.) which is a primary feature of improvisational creative collaboration.

The overall aim of this chapter is to show how an enactive approach to computational creativity can make it easier to think about, design, and build creative computers, especially those that are able to improvise in real time collaboration with human users. To situate and motivate our contribution, we first describe the traditional approach and cognitive assumptions popular in the field of computational creativity. Next, we introduce the cognitive science theory of enaction and describe creativity through its theoretical lens. Then, we present our enactive model of creativity and explain how it can help operationalize the theoretical premise of enaction for use in co-creative systems.

3.3 Theoretical Premise of Computational Creativity

Many examples and perspectives of computational creativity were presented in the related work of Chapter 2, and here we will describe the field from a high level to extrapolate a general view of the cognitive assumptions and motivations. Bodens [11] work is perhaps the most foundational and far reaching. It generally describes creativity from the perspective of the information processing (IP) paradigm of cognitive science, accounting for creativity as a largely search-based computational process. The information processing paradigm in cognitive science describes cognition as the formation and manipulation of representations and knowledge structures in the brain of an agent. Based on the IP account of creativity provided by Boden, important research questions for computational creativity include methods of acquiring knowledge as well as different procedures for combining, manipulating, and transforming that knowledge to produce novel and valuable products [11]. This theoretical approach is popular in the field and has been successfully applied to generate novel works of art in a variety of creative domains, such as narrative [20,129,131,172], poetry [122,127] , ideation [1], games [26, 76, 139], analogy [58, 73], and creative design [102, 149].

A popular software architecture for generative computational creativity is a type of handcrafted knowledge-based system. For example, first the systems first reads or interprets a large corpus of material into structured representations that it uses as its knowledge base. To make the systems more creative, the corpus can be carefully selected to lead to more interesting combinations, such as twitter posts and news articles [23, 173]. These representations form the conceptual space the agent traverses to find interesting combinations to produce novel output [11]. For example, a poetry generating system might parse a news article into structured representations that can then be spliced and recombined according to hard coded rules of poetry (meter constraints, rhyming patterns, etc.). The conceptual space itself can be restructured to reveal additional mappings and traversals within it, which is called transformational creativity [11]. Finally, those spaces are systematically traversed to piece together a novel creative product, which is outputted to the user. These types of creative systems typically yield bounded and discrete creative artifacts as their output.

While generative systems certainly produce results that are compelling, it is another question entirely to consider whether the process by which they were created reflects the human creative process. For this reason, the algorithms and cognitive architectures of creative systems have also been extensively evaluated [22, 179]. The critical factor here is that the metrics used for determining what is creative is inherently based on theories of creativity and cognition. Using the information processing paradigm yields one set of metrics related to the way in which information was processed in a system, while using enaction results in a different set of evaluation metrics focused more around the interaction of an agent with the environment in which it is embedded. Creativity, from an information processing perspective, need not necessarily occur in conjunction with a body embedded in an environment. The 'creativity' that generative systems exhibit can occur in an abstracted manipulation of symbols without a perception-action feedback loop with the environment. While the end product of generative systems may resemble something we might expect of a creative human, we argue these systems leave out one of the most fundamental ingredients to human cognition and creativity–the interaction of a creative agent with the environment, the feedback loop that emerges to guide that interaction, and knowledge constructed through these experiences over time.

Creativity is inherently interwoven into human intelligence and displayed in a myriad of contexts and conditions. As a result there are many different perspectives and lenses through which creativity has been analyzed and studied. In particular, research has focused largely on four dimensions of creativity: the person, product, process, and press or environment. Much progress has been made in each of these threads. The theory of everyday creativity proposes a views that cuts across these four pillars of creativity research to propose a continuum of creative cognition ranging from everyday actions to extraordinary acts of creative genius [142, 145]. The primary contention for everyday creativity is that there is a common underlying cognitive process across many different creative activities [145] (c.f. [118, 121]) for a description of similar description of this continuum hypothesis in scientific creativity and reasoning). Similarly, we suggest it may be more effective to investigate creativity as an emergent product of more fundamental cognitive processes operating in unique ways.

The novel theory of creativity proposed here is focused around sense-making, the process by which a cognitive system interacts in an environment to gradually build meaning through emergent perception and action feedback loops and prior experience. To provide a full view of our proposed theory of creativity, we begin by sketching out the enactive perspective and the role that sense-making plays in this view. We first give a brief history to introduce the enactive paradigm and briefly describe its main concepts. Then, we describe the prominence of perception and action for describing cognition as an emergent property of autonomous agents interacting with the environment. Next, we examine the role of goals and planning in the enactive perspective. Then, we describe the implications this approach has for creativity research and propose a novel view of creativity leveraging sense-making as the fundamental cognitive process responsible for creative cognition. A novel enactive model of creativity is presented that begins to systematize and formalize the conceptual framework of creative sense-making described here. Finally, we consider multiple domains of creativity from the perspective of creative sense-making to show its utility for describing and understanding the creative process.

3.4 Introduction to Enactive Cognitive Science

The term *enaction* was first coined by Varela in the book The Embodied Mind [170]. While the ideas of embodiment advocated by Varella have become largely accepted in the cognitive science, AI, and robotics communities, the enactive paradigm within which Varella presented embodiment is only beginning to be disseminated into AI, mostly with respect to developmental robotics, i.e. designing robots that learn through their own embodied experience [65, 174, 176]. Enaction is gradually gaining in popularity since the publication of Stewart's et al. foundational book *Enaction: Toward A New Paradigm for Cognitive Science* [162]. Evidence of its growing influence in cognitive science can be seen in the proliferation of cognitive science articles utilizing the premise of sense-making as well textbooks beginning to formalize the novel paradigm, such as *Artificial Cognitive Systems: A Primer*, which recruits many of the main concepts of enaction for understanding how to design and implement embodied agents [175].

The fundamental basis of the enactive paradigm is that cognitive systems gradually develop their own understanding and knowledge of the world through their interaction with the environment [174]. The process of developing that understanding is in service to sustaining the life form. Cognition, from this perspective, is the "process whereby an autonomous self-governing system acts effectively in the world in which it is embedded" [174]. The main divergence from information processing accounts of cognition is that cognition is seen an emergent process that cannot be separated from an agent embedded in and interacting with its world. The world does not merely serve as inputs to cognition, but the environment is fundamentally constitutive of cognition through the process of interaction. In particular, enaction emphasizes the role that perception plays in guiding and facilitating emergent actions [38]. A short definition of enactivism by Havelange [82] will help summarize this distinction:

> "Here, cognition is no longer considered as a linear input/output sequence (as was the case in classical cognitivism [i.e. information processing accounts of cognition]), but rather in terms of a dynamic sensorimotor loop by taking into account the fact that actions themselves produce feedback effects on subsequent sensations. Action is thus no longer a simple output; it becomes actually constitutive of perception. What is perceived and recognized in perception are the invariants of the sensorimotor loops, which are inseparable from the actions of the subject."

There are five core ideas to the cognitive science theory of enaction: autonomy, sense-making, emergence, embodiment, and experience. Together, these concepts weave an alternative paradigm of cognition that is focused largely around the interaction of an agent with the environment in which it is embedded.

3.4.1 Autonomy

The concept of autonomy refers to how cognitive agents operate independently based on their own intrinsic laws to sustain themselves and continually generate their identity through interaction with the environment according to those laws. A system is defined as autonomous when it is able to define and change the laws that govern its interaction with the world by casting a web of significance (i.e. meaning) on the elements in its environment as they relate to sustaining and generating its autonomous identity [170].

3.4.2 Sense-Making

As autonomous agents interact with their environment, they gradually detect patterns of regularities in their interactions, which help them understand the environment in a process referred to as sense-making [39]. When a particular type of action produces a predictable result, the agent can begin to form expectations for their actions and determine regularities in their environment that gradually begin to make the environment more meaningful. These percept-action pairings have been termed *sensorimotor contingencies* meaning that certain types of motor actions create a predictable response that is perceived through the senses [124]. Once an agent knows approximately what to expect from performing an action in a particular situation, that agent has made sense of that element in the world. The key point here is that this process of building meaning has to be done interactively, through the agents own actions and perceptual system creating these contingencies that help the agent predict how its environment will respond to its actions.

3.4.3 Emergence

Emergence is critically important to enaction because enactivists claim meaning emerges through interaction between an agent (or multiple agents) and an environment. This feature of enactive cognition relates to the research on situated cognition [164, 165], which describes how humans don't merely execute plans to accomplish actions in the world, but rather rely upon information in the environmen to inform action, such as object affordances that serve to guide interactions without formal plans or representations about that interaction the brain of the agent. Situated action emphasizes procedural and sensorimotor knowledge that is enacted in a continuous environment through interaction. The idea of emergence in enaction shows how these types of situated and continuous interactions lead to emergent meaning in the moment, which could be highly influenced by other agents in that environment.

As a result of viewing cognition as an emergent phenomena, it cannot be studied in isolation from the context and environment within which an agent is embedded. Cognition is what happens in the course of an agent making sense of its environment [40].

3.4.4 Embodiment

Enaction also adheres to the embodiment theory in asserting that since agents must act with their body to make sense of the environment, their bodies inherently constrain and afford certain types of interaction. Given that interaction is central to the paradigm of enaction, the body that does the interacting plays a critical role in the formation of knowledge and the relation between perception and action.

In the enactive paradigm, perception is not a passive reception and classification of sensory data in structured representations, but rather an active process of visually reaching out into the environment to understand how objects can be manipulated [72, 124]. Here, cognition is seen as a process of anticipation, assimilation, and adaptation, all of which are embedded in and contributing to a continuous process of perception and action. This type of enactive perception minimally involves a negotiation among the following factors: 1) The agents intentional state; 2) The skills and bodily capabilities of the agent; and 3) Perceptually available features of the environment that afford different actions such as size, shape, and weight (e.g. is it graspable, liftable, draggable, etc. as elaborated by Norman in [126]).

Sensory data enters the cognitive system and irrelevant data is suppressed and filtered [68]. Objects and details of the environment that relate to the subject's intentional goals appear to conscious perception as affordances, which can grab, direct, and guide attention and action [126]. Each time the individual physically moves through or acts upon the environment, that action changes the perceptually available features of the environment, which can reveal new relationships and opportunities for interaction.

Instead of being purely the result of symbolic manipulations occurring in the brain, for enactivist researchers, cognition is the process by which meaning emerges through interaction with the environment. Without an environment to make sense of, the brain and the body would have no purpose or direction. In this way, perception is said to be for action, i.e. perception has a purpose and reason for existing in relation to action [125]. The process of perception is not an independent module or cognitive mechanism, but rather it is critically tied to action [125]. Perception informs action and taking actions reveals additional percepts that further inform the agent what type of actions are possible and relevant. Through this dynamic process of interaction, cognition emerges as a process by which the agent determines how to effectively manage and engage with the environment.

3.4.5 Experience

Finally, the experience of the agent is shaped by the meaning it has imbued into its environment and this influences how future actions are generated and evaluated. While the body constrains how the agent might interact with the world, the meaning it generates and the knowledge an agent acquires as a result of that meaning is always based on experience. The enactive approach takes first person experience and awareness of the cognitive agent as the starting point. It advocates for an intelligent perception and action system that pairs interesting actions and related percepts as a coupling that are stored to guide future interactions. Enaction is rooted in the notion that cognitive agents always experience reality as a continuous interaction with the world and any investigation or model should have interaction as its fundamental constituent. As a result, perception, action, and goal formulation are modeled much differently in the enactive paradigm than in the information processing paradigm.

3.4.6 Goals and Directives

From an enactive perspective, intelligence and creativity involve knowing how to change the flow of sensory information in order to explore possibilities for action, i.e. leaning in closer to get a better look at something. It is often simply easier to act on the environment and experiment with how different interactions affect the system than representing it in its entirety and performing symbolic processing on those representations like the information processing perspective proposes [124]. Even at the level of social interaction with an intelligent agent, an enactive approach tries to avoid postulating high-level cognitive mechanisms at the core of our intersubjective skills. The co-evolution of a communicative/creative process is seen here as a gradual unfolding in real time of a dynamic system spanning a human subject, the environment, and agents within it. In this view, intentions emerge but are also transformed in and through the interaction with other agents and the environment.

Thus, instead of describing creative behavior as goal-based planning and information processing, we have adopted the enactive terminology of directives [52]. A *directive* is a loose intention that directly influences what things appear interesting or salient in the environment, and how specific types of interactions might provide more information about emerging hypotheses. A directive is similar to a goal in that it can be reflected on, elaborated, and specified in more detail, but it is critically different from the current notion of goal in planning-based AI because it does not constitute action in any way. A directive constrains and suggests potential actions that could yield productive changes in an emergent process of sensemaking. See Figure 8 for an illustration of goals compared to directives. To summarize the idea of a directive, a directive does not dictate action; it selects a filter for perception that (we propose) enables a perception-based reasoning process we call perceptual logic. Actions are executed in response to situated features of the environment. Some actions are executed in service of tasks (i.e. executing a plan), while other actions help gain different perspectives, including changing physical location as well as changing the directive with which a scene is analyzed. These changes in perception serve to provide more sensory information and resources with which to help the agent reason and determine what an appropriate next step would be. This process is guided by attention and the awareness of the agent and is inherently based on the temporal flow of experience and the dynamics of interaction with the environment.

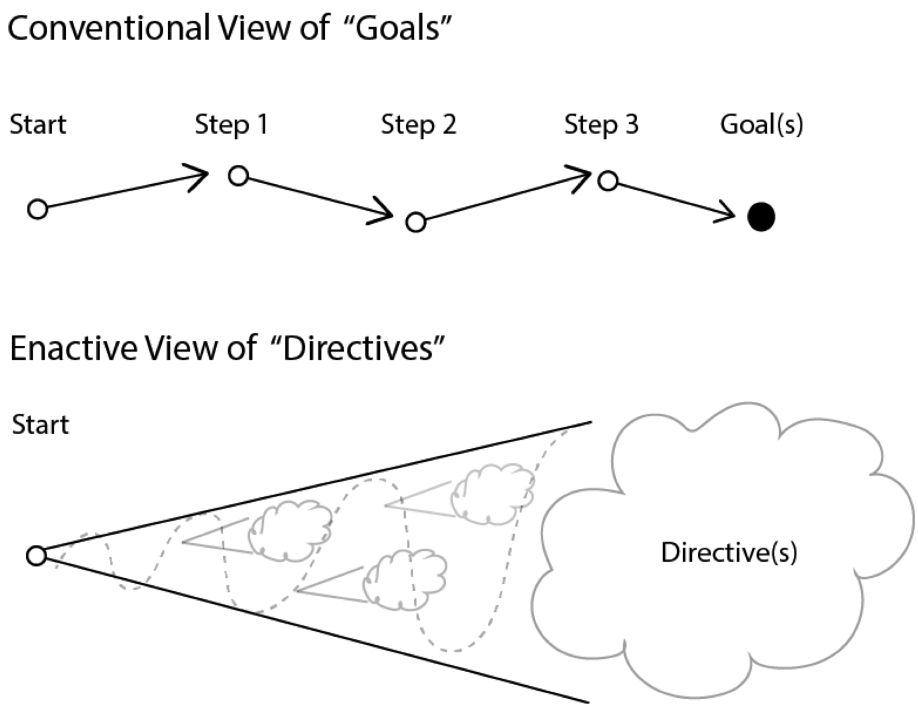


Figure 8: Comparing goals and directives. Plans are usually linear with a series of steps toward a specified end-state whereas directives are vague and gradually refined through a process of interacting with the environment and defining tasks that explore the problem space outlined by the directive.

3.5 Examples of Enactive Creativity in Multiple Creative Domains

The literature on creativity provides evidence supporting the enactive approach to creativity with research on *thinking by doing.* There is a multitude of evidence demonstrating how both representational and non-representational artists plan their artworks using sketches, studies, and other ways to simulate artistic alternatives [104]. Sketching reduces cognitive load and facilitates perceptually based reasoning [148]. In many creative domains, individuals generate vague ideas and then use some form of sketch or prototyping activities to creatively explore, evaluate, and refine artistic intentions [31]. Sketching allows creative individuals to think by doing. When an action or idea is materialized in some way, the perceptual system is rewarded with richer data than pure mental simulations and abstract reasoning. Additionally, cognitive resources that would have been used to simulate the action (i.e. consciously visualizing the situation) are now freed for other tasks such as interpretation and analysis [153].

3.5.0.1 Architectural Design

One obvious example of using sketch to think by doing can be found in the task of planning the spatial configurations in the architectural design process. As addressed above, generating an entire artifact with all of its details directly from the mind is virtually impossible for a designer [148]. Instead, designers use sketch to facilitate their thought process. For this reason, Schon refers to design sketching as a reflective activity that prompts new ideas and facilitates design creativity. Designers can rapidly explore and alter sketches in real time to help evaluate concepts and explore ideas [167]. When starting the design process, designers choose different materials, tools, and media to present the initial ideas from their minds to explore the constraints of their problem [8, 148].

When designers interact with their tools, they might need to adjust their actions in order to achieve their needs. For instance, when drawing a sketch to study the form of objects and spaces, they may need to constantly adjust the 'next steps' in order to solve the design constraints, such as not enough space, too long, too much curvature, etc. In this way, sketching enables designers to offload part of their process onto the environment in order to facilitate thinking. The designer does not have to consciously visualize the design once they sketch some aspect out, which frees up cognitive resources for reflecting and analyzing the ideas [46, 75, 167, 169]

Figure 9 illustrates a typical spatial plan of a student center in a bubble diagram. Since the plan entails many spaces, the designers would have to write down all the space names so that related spaces are located next to each other. They would also use arrows to represent the main circulation paths between two spaces. Each time a new space is added or an arrow is inserted, the designers flow of sensory information changes and they might discover new problems or opportunities that were not apparent before [167]. Sketching facilitates their creativity and reasoning process through a dynamic perception-action feedback loop whereby new meanings are gradually constructed through a negotiation with the design materials (i.e. sketch, physical models, computational models, etc.). Design sketching enables a type of *thinking by doing* that can help facilitate the creative process [47].

Experienced designers also change the granularity of their perception to reason about sketches at different levels. When focusing on individual details, an architect might imagine how a particular corridor might feel to walk through. Then, they could shift to a global perspective that considers the overall theme and consistency of the whole building design [75, 167].

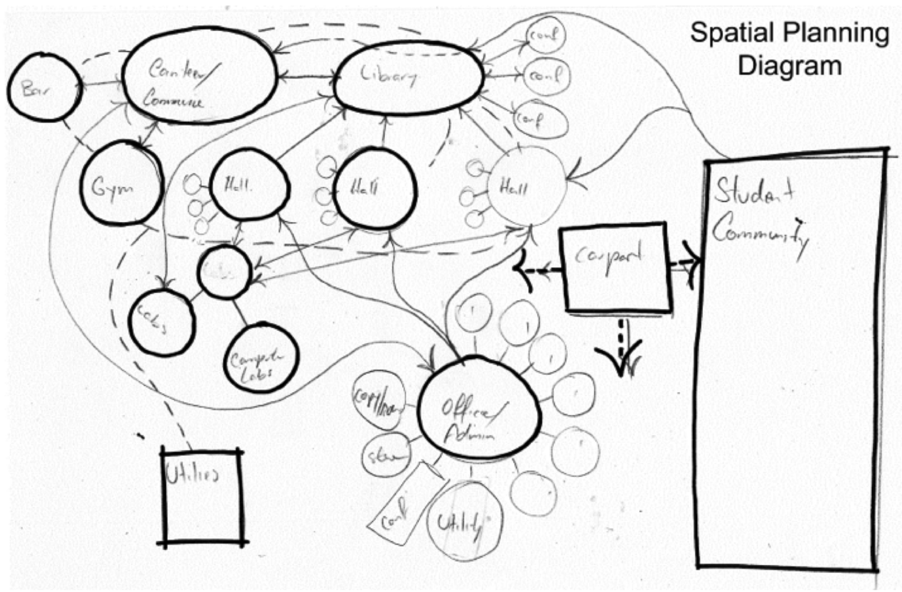


Figure 9: Spatial layout of a school student center design (courtesy of Kyle Doggett).

3.5.0.2 Musical Performance

The enactive nature of creativity can also be seen in live musical performance. A classical musician, for example a trumpet player, will need to feel the acoustic effects in a concert hall before his performances. For instance, he may extend the ending of a sound in a concert hall that has a dry acoustic effect. We propose the expert trumpet player has a well established set of percept-action pairings (creating his expert perceptual logic) that have to be tuned to the particular performance space because the actions he will take in the performance will result in a slightly different perceptual feedback process than his normal practice space. Thus, he has to actively feel and explore the sounds of the space to align his perceptual logic with the specifics of the exact situation. Furthermore, during performance, he will also listen to the mixture of his trumpet sound with other sounds to make real time adjustments to achieve the desired general effective (i.e. the directive, such as playing a sad tune).

3.5.0.3 Visual Art

When an agent decides to embark on a creative activity in and of itself, we can again use sense-making and fluctuating cognitive states to describe creativity within this process. In the case of improvisational art, the goal is not fully developed before the artwork begins. Thus, the agent's cognition begins in an unclamped state during which an open-ended sense-making process is utilized whereby potential goals as well as the ways of achieving them are explored. As ideas emerge through interaction, cognition begins to become clamped on each individual task as well as an overarching directive or artistic theme. The artmaking activity would contain many sense-making cycles that modulate between clamped and unclamped cognition, essentially employing different ways of seeing, evaluating, and interpreting the artwork through time. Different artists certainly have different processes as far as how their clamping/unclamping process unfolds. Some improvisational artists may begin with a single idea and develop it in-depth until it leads to the next idea, whereas other artists may start many different ideas on the canvas (each with its own sense-making process leading to a clamped state of cognition and coupled interaction) and then visually analyze how they interact with each other.

With representational art, the process would look different. Since the artist has a clear goal of the end state, there will be longer periods of unclamped cognition early on to determine a general approach to achieve the goal. Once the artist clamps on a particular course of action, she will begin accomplishing tasks that are clearly outlined from the onset. There will be less fluctuation in the unclamped and clamped cycles of cognition in representational art since the plan is relatively static from an early stage. In this process, sense-making would largely occur during the process of comparing the relative balance of each region as they progress together to form the overall compositional goal.

The enactive nature of creativity in visual art is demonstrated well by findings showing that expert artists often step away from their paintings to gain a new perspective [185]. Yokochi [185] analyzed the painting process of a famous Japanese painter. He found that the artist began with a vague directive that is then refined and explored through interacting with the painting. Each new line adds additional constraints and affects all the existing constraints created by previous lines. Once the painter takes a step back to understand her last contribution in terms of the overall picture, she may find that her last contribution actually disrupted the overall balance of the piece. Although she doesnt have a specific end-state for the painting in mind, one of the directives guiding her work may relate to achieving an overall balance in the composition. This directive does not determine what contributions to make, but it helps point out inconsistencies and visual tensions that need to be addressed to achieve balance.

Let us suppose that the artist found 5 areas of the drawing that all violated his sense of balance due to his last contribution. He would then select one of those areas and defines specific painting tasks that he predicts will help achieve balance. Once the first of those 5 areas is complete, the artist could take another step back and realize that his latest contribution makes the left side of the artwork look kind of like a face, which they may like and find promising as an idea to pursue further. The artist might then update their overall directive to creating some kind of abstract face. Once this directive is adopted, the entire canvas is analyzed with respect to face-like features. Given this new constraint, the artist sees additional opportunities to change the drawing and would then select specific painting tasks that contribute toward the current directive. Here, the directive is dynamic and always evolving through interaction with the environment. The feedback offered by actually producing a change in the environment spurs new ideas and interpretations that can change the overall directive. The directive determines the constraints and affordances that are consciously available to the painter's perceptual processes.

Attention of the agent drives the system by changing the flow of sensory information. Depending on the current directive, the system perceives sensory information in different ways. At this point, the reader might ask: how can the same sensory information be perceived in different ways? If we imagine sensory input as a flow through time, we can then consider adding different *lenses* to perception to filter that sensory input in different ways. Different filters make different features of the environment salient. If they are salient enough, they will demand the attention of the individual, which might then prompt subsequent interaction. The directive guides attention towards facets of the environment that are relevant to the current intention of the agent. The old adage *when you have a hammer everything looks like a nail* is quite illuminating to consider in this context. Once a hammer is picked up, the general directive of hammering is established, and this directive guides attention and action, which results in things being perceived in terms of their *hammerabilitiy.*

The experiential knowledge of expert artists is rooted in acquiring percept-action pairings about how altering the flow of sensory information can reveal additional possibilities for action. The percept-action coupling, in this case, relates to how moving the body helps to reveal different viewpoints and visual relationships. There is no preset specific goal driving the artists decision to step back, and there is not a *step-back-and-think script* that the artist executed at predefined times. Instead, there might be some open questions about how to interact with different regions of the artwork and a vague intention to address those concerns. Stepping back helps artists unclamp from local tasks to think about how interacting with many different ideas and areas might affect the overall vague intention. The creative behavior of stepping back is actually an emergent by-product of how cognition and creativity work. The fact that the artist stepped back (her behavior) is therefore not as important as why she stepped back, i.e. how she knew that stepping back was the right thing to do. An expert is an expert precisely because she knows how to direct her attention and manipulate the flow of sensory information through interactions with the environment to explore and evaluate possibilities for further action.

Ultimately, it is the continuous perception-action feedback loop that actually determines actions. Instead of thinking of action as a series of behaviors executed like scripts or plans, we can think of action as a continuous improvisation with the environment guided and evaluated with respect to the currently active directive. Attention and the conscious experience of the agent becomes the common thread that stitches the flow of each individual action together.

3.6 Collaborative Creativity

This proposed theoretical framework also helps to shed light on creative collaboration. When two people collaborate, their states of cognition and transitions between them significantly impacts the nature of the collaboration. For example, when two individuals are drawing together and they begin immediately to work on independent sections on different activities, they are independently clamped. When one person leads the interaction and determines most of the new structural elements that are contributed, their partner can be said to be clamped onto their creative trajectory. Conversely, two people can work together to define the general structure and strategy for interaction throughout the collaboration in such a way that both parties mutually contribute to the shared meaning structure that is used to anchor their clamped states of cognition in a co-regulated mutual coupling. Additionally, one can intentionally try to disrupt a collaborator from their task to help them explore new ideas. For example, collaborators can make contributions that surprises their partner, which would cause a temporary tensions followed by resolutions that might deepen the meaning of the collaboration. See Chapter 6 for more details on participatory sense-making during improvisational collaboration in open-ended creative domains.

3.7 Conclusions

Computational creativity has the potential to radically change what it means to interact with computers. However, in order to reach its full potential, the field needs a cognitive theory of creativity that accounts for the enactive nature of creativity, including improvisation, collaboration, and a tight feedback loop with the environment. In this chapter, we provided a brief summary of the current state of computational creativity and pointed out the shortcomings of the traditional information processing view of cognition. We argued that the new cognitive science paradigm of enaction provides a helpful way to reframe creativity and potentially solve some of the long-standing hard problems that both artificial intelligence and computational creativity face. The theory of enaction was used to describe creativity in design, music, and visual art to show its potential for generalizability and descriptive power. We also presented the enactive model of creativity that formalized the enaction theory in a computational model. The primary design principle of the enactive model of creativity is to design interactions like a conversation where each party tries to make sense of contributions and respond appropriately given the history of interaction.

CHAPTER IV

DRAWING APPRENTICE SYSTEM DESIGN

4.1 Summary

This chapter describes the technical details of the Drawing Apprentice prototype, which is a web-based co-creative drawing partner built to serve as a technical probe to explore what types of interaction designs and machine learning approaches might facilitate particpatory sense-making. The chapter begins by summarizing the design rationale for the system and providing a overview of the user interface and user experience. Next, the drawing algorithms are described, including early reactive algorithms that tranform user input as well as later algorithms that incorporate object recognition and object drawing. Technical challenges for recognizing sketched objects in open-ended contexts are described as well as our proposed solutions to these challenges, such as grouping lines for classification and placing objects on a canvas.

4.2 Introduction

The Drawing Apprentice is a co-creative drawing agent that analyzes the users input and responds with artistic contributions of its own on a shared canvas. It is a technical probe to explore what type of interaction design and machine learning approaches facilitate participatory sense-making in open-ended improvisational domains like collaborative drawing. Inspired by the theory of enaction, the system is designed to engage in participatory sense-making by coordinating its actions through real-time feedback.

In the domain of drawing, there are a number of critical variables for a co-creative agent to consider when devising its sketch contribution, such as what to draw, when it should be drawn, where to place it on the canvas, how it should be drawn (i.e. the manner of drawing the lines), and why the agent should draw a particular element. Simultaneously determining what the correct answer is for these variables in an open-ended collaboration presents a monumental challenge for a creative agent given the nearly infinite variety and dynamism of the artistic intentions of users throughout their creative process. To delimit the creative responsibilities of the agent and help facilitate a coordinated collaboration, our approach offloads some of the creative decision making processes onto the user through direct interface controls and feedback mechanisms. This hybrid approach enables the user to maintain a degree of control over the agent's drawing activities while still affording creative and unexpected contributions within defined boundaries.

One of the primary technical questions we have asked in developing this prototype is what the system should be focused on learning and how it should acquire the proper knowledge to sustain a drawing collaboration. Early versions of the system [33,35,36] focused on imitating the users sketch input using a variety of transformation algorithms. User feedback was employed to train the system about the users stylistic preferences. This approach was devised primarily for use in abstract drawing to provide users with novel input to stimulate ideas and help explore and transform the creative space of the drawing in unexpected ways. However, findings from user studies (described in the next chapter) indicated that additional factors need to be considered when users attempt to draw representational objects formed by multiple lines with a clear artistic intention.

User study findings (described in Chapter 7) showed a few key requirements for co-creative drawing agents meant to facilitate participatory sense-making. There were several critically important factors that contributed to whether the agents contribution made sense to the user, such as demonstrating spatial awareness while drawing, drawing visually similar elements, as well as aligning with the prevailing perceptual logic of the users contribution [35]. Spatial awareness refers to the agent not drawing over existing shapes, and drawing in a similar region as the user. Visual similarity describes how users could more easily understand contributions that were clearly visually related to their own. Users also expected the system to understand what they were drawing, including the logic behind the application of patterns they were producing and objects they were drawing. Participants reported wanting the system to contribute to their drawing in a way that reflected some understanding of what they were drawing. For this reason, we concluded object recognition is an important skill for a co-creative drawing agent. Enabling the agent to understand what types of objects are being drawn opens up the opportunity for more coordinated collaboration through shared understanding.

We devised an additional machine learning module emphasizing understanding the users drawing and responding with appropriate objects by employing object recognition in order to increase the agents ability to build shared meaning through interaction. Implementing object recognition in a real time and open-ended context presents significant challenges both in terms of how to structure the sketch input data for classification as well as ensuring a high accuracy in a short amount of time to facilitate real time interaction. The remainder of this chapter describes the important technical components of the Drawing Apprentice system architecture, and it explains how they relate to the user experience.

4.3 System Overview

At the highest level, the Drawing Apprentice is a co-creative agent that analyzes user input lines in a virtual canvas, responds to those lines using a variety of algorithms, and outputs new lines onto the same canvas (see Figure 2). The system is implemented as a web application (see http://adam.cc.gatech.edu/DrawingApprentice/) with a client-server architecture that enables multiple people to collaborate with each other as well as the agent from the Drawing Apprentice system. It was designed for use with stylus- or touch-based interactions, but a mouse can also be used.

The user interface of the Drawing Apprentice system has three main components: a palette of functions to control the agent, a palette of drawing tools, and a shared canvas. The agent palette contains buttons for controlling the five drawing modes (section A of Figure 10), voting buttons for providing feedback (section B of Figure 10), and the home base of the character icon representing the co-creative agent (section C of Figure 10). The agents interpretations of the users drawing activities appear as a speech bubble in this home-base. The drawing palette consists of functions traditionally associated with drawing applications, such as selecting a color, line thickness, saving the image, and starting a new canvas.

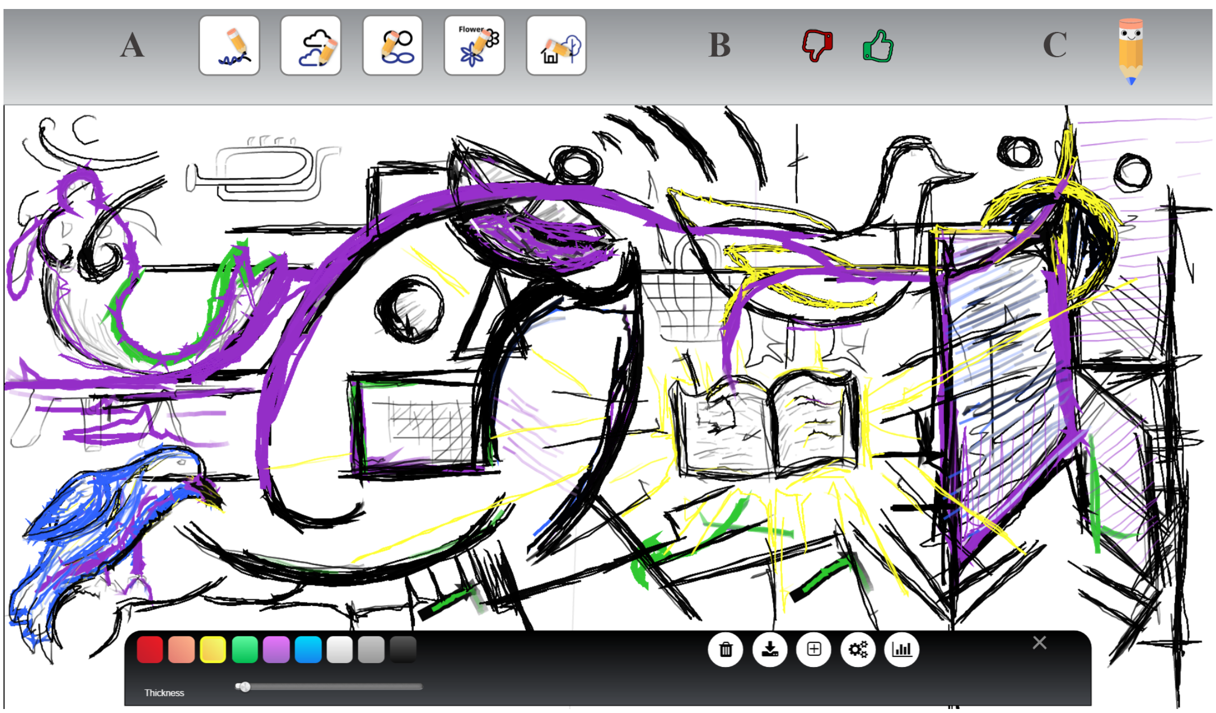


Figure 10: The Drawing Apprentice Interface and Example Drawing. Top panel offers the communication channel between the user and agent. Bottom panel contains conventional drawing functions.

After the agent finishes its turn, the user can provide feedback via voting buttons to inform the agent whether the user liked its contribution. This voting information is used to learn the aesthetic preferences of each user and fine tune what types of contributions it makes within each of the drawing modes.

4.3.1 User Experience

Instead of hard-coding knowledge into the agent to make it creative, the Drawing Apprentice is designed to extract critical information from a user that is already exhibiting intelligence and creativity given the target domain in an open-ended creative task. Therefore, the most general design principle is whenever possible, offload any higher-level cognitive tasks to the user, i.e. enable the user to manually specify the boundaries within which the agent should operate throughout the interaction. The agent does not need to know why those parameters exist if those values help facilitate effective collaboration. Known as the Eliza effect, researchers have demonstrated that users attribute intentionality to virtual agents if they appear to understand the context of the situation, even when the system may not understand the actions it performs [177, 178]. Through interaction and user experience design, defining constraints and parameters can be smoothly integrated into their creative flow or removed altogether as the machine learning algorithms grow in complexity and sophistication.

4.3.2 Turn Taking

Turn taking was designed to facilitate emergent interaction dynamics, meaning the number and length of the agents lines are dependent upon the users recent contributions. As soon as the user ends their current line, a timer begins. If this timer passes the arbitrary value of 2 seconds before the user begins their next line, their turn ends, and the system starts to draw. The systems turn will be approximately the same number of lines as the users turn to mirror the interaction. However, the user may begin drawing at any point, which can lead to synchronous collaboration.

4.3.3 User Feedback

During human-human collaboration, artists are able to leverage extremely subtle and implicit cues to facilitate coordination during joint activity. To begin approximating this feedback-based coordination described in the enactive literature, we enabled basic feedback mechanisms, such as binary voting, to explore what types of interaction mechanisms may be effective in facilitating coordination and participatory sense-making through real-time feedback. From the perspective of optimizing the machine learning algorithms we have employed, the human should be required to provide feedback on every contribution made by the system. However, human users are never required to provide feedback in order to maintain the creative flow of the artistic experience. Instead, users are given the choice to vote whenever it occurs to them, which can be highly variable between users. This presents an interesting opportunity where improving the user experience design might potentially improve the performance of the machine learning algorithms (since more feedback helps train the system).

4.3.4 Character Design

To simulate the dynamism and embodied nature of real-time human collaboration, the Drawing Apprentice character draws lines dynamically, meaning lines do not appear at once in full, but are gradually animated through time until their completion. Dynamic line drawing is meant to provide a sense that the system is going through the embodied act of creating a line.

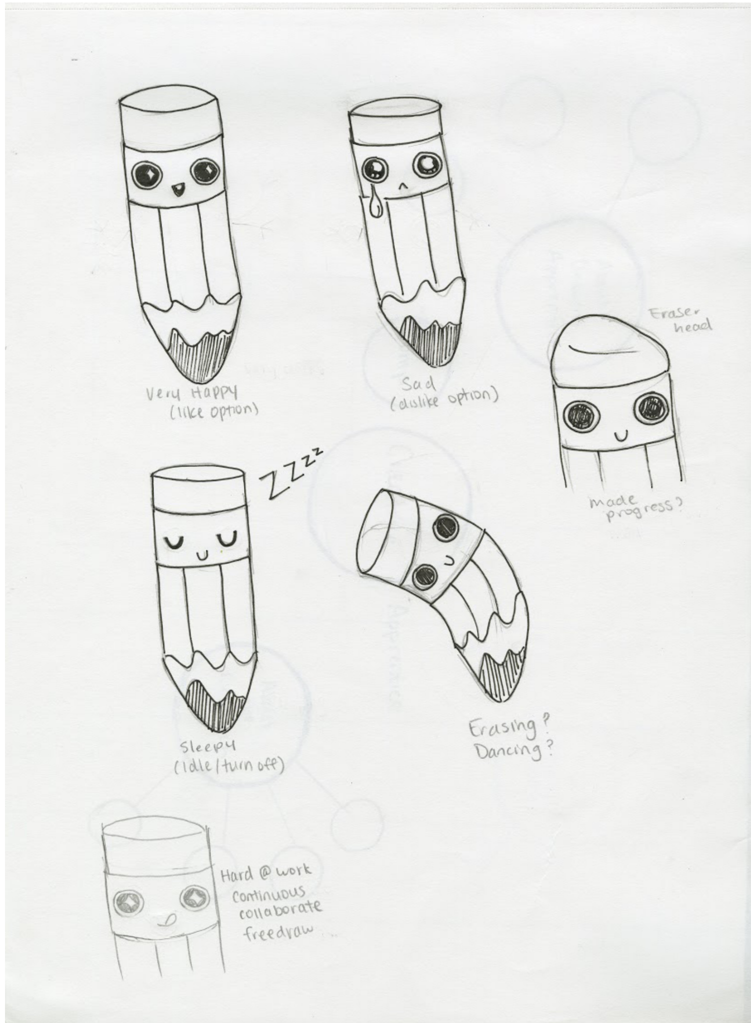


Figure 11: Sketches exploring character expressing affect as a means of feedback (designed by Lisa Li)

Pilot studies revealed the importance of having a character represent the Drawing Apprentice on the canvas, which is a sentiment echoed in the literature on embodied virtual agents [6, 80]. In an early version of the prototype, the character only appeared while the system was drawing. In the pilot studies and during demos watching and talking about the character seemed to excite the participants. Multiple users requested a permanent presence, or home for the character while it was not drawing. We improved the character design and created a home base for the agent to return to after it has completed its drawing. In our context, character design may be able to improve the perceived performance of the agent. With the right animations and visual design, the system might appear more creative and intelligent without any change to the algorithms. For example, When a drawing mode is selected, the image on the button animates to provide a prototypical demonstration of what the drawing mode entails to help the user understand what to expect from the system. Figure 12 demonstrates that animation depicting 5/25 frames demonstrating the hypothetical input and response for a particular drawing behavior.

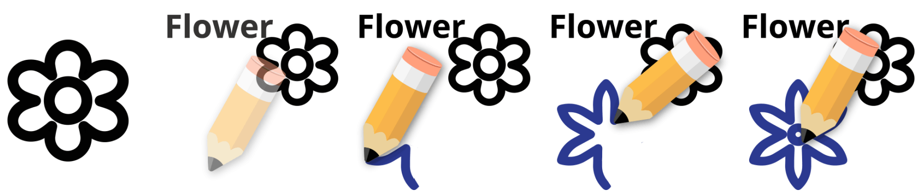


Figure 12: Button Animation Demonstrating Drawing Mode Response (designed by Lisa Li)

To increase the shared knowledge between the user and agent, we implemented a speech bubble animation. In this speech bubble, the agent tells the user what it thinks they are drawing as well as what it plans to draw when the object recognition drawing modes are selected.

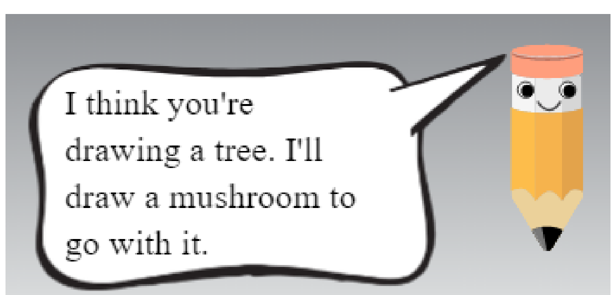


Figure 13: Agent speech bubble to communicate the agents interpretation of the drawn object and what it will draw next.

4.4 Drawing Algorithms

The drawing modes can be divided into two functional categories: responding to individual lines and responding to groups of lines. The first three reactive drawing modes are tracing, transformation, and mimicry. In tracing mode, the system sketches over the input lines to provide texture and depth to existing lines. In transformation mode, the system manipulates the input lines by either rotating, scaling, or translating them. Finally, the mimicry algorithm skews the initial input by stretching and shrinking various aspects of the input line. This category of drawing modes was designed primarily for use in abstract drawing to provide the user with novel input to stimulate ideas and help explore and transform the creative space of the drawing in unexpected ways. While these drawing modes can provide interesting contributions for abstract collaborative drawing, findings from user studies (described in the next chapter and [35]) indicated that additional factors need to be considered when users attempt to draw representational objects formed by multiple lines with a clear artistic intention.

To enable the system to collaborate with representational contributions, we created the second category of drawing modes that groups input lines and attempts to classify the type of object the user is drawing, as illustrated in the red block of Figure 2. First, the system must determine which lines to group (see Line Grouping section), then an image is formed from those lines and sent to a precomputed Convolutional Neural Network model for classification. Finally, the system employs one of the grouped-line drawing algorithms and outputs the results to the shared canvas.

4.4.1 Reactive Algorithms

The first category of drawing algorithms react to individual line input from the user. Leveraging user input as a source of creative contributions is meant to help build meaning by modifying the users own intention through manipulating their lines. In early prototypes, these algorithms were assigned to a spectrum of creativity controlled by the user with a creativity slider on the user interface, as shown in Figure 14.

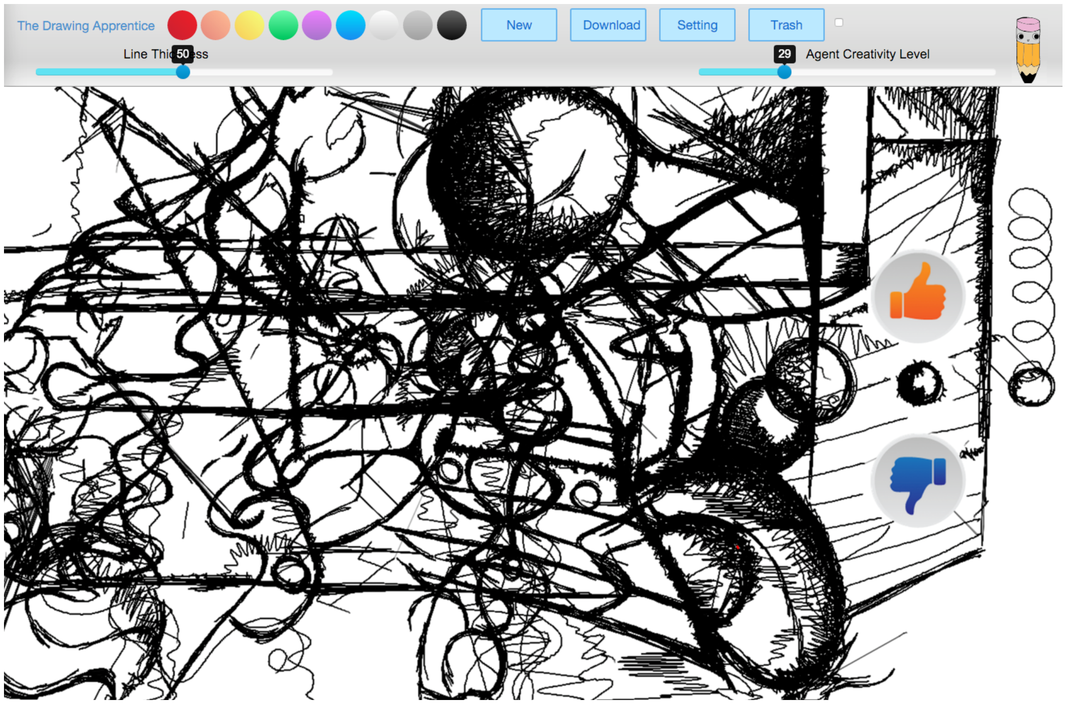


Figure 14: Early interface with creativity slider controlling agent drawing behavior and example drawing with reactive algorithms

This creativity slider was later replaced with discrete state buttons as users found these more intuitive and easy to understand (see Figure 10 to see new interface design). The buttons and their features will be described in later sections. Here, we describe the initial slider implementation and the technical details behind it in order to provide context for the user studies described in the next chapter that utilized the slider prototype. The system diagram in Figure 15 shows how the early prototype functioned with the reactive drawing algorithms.

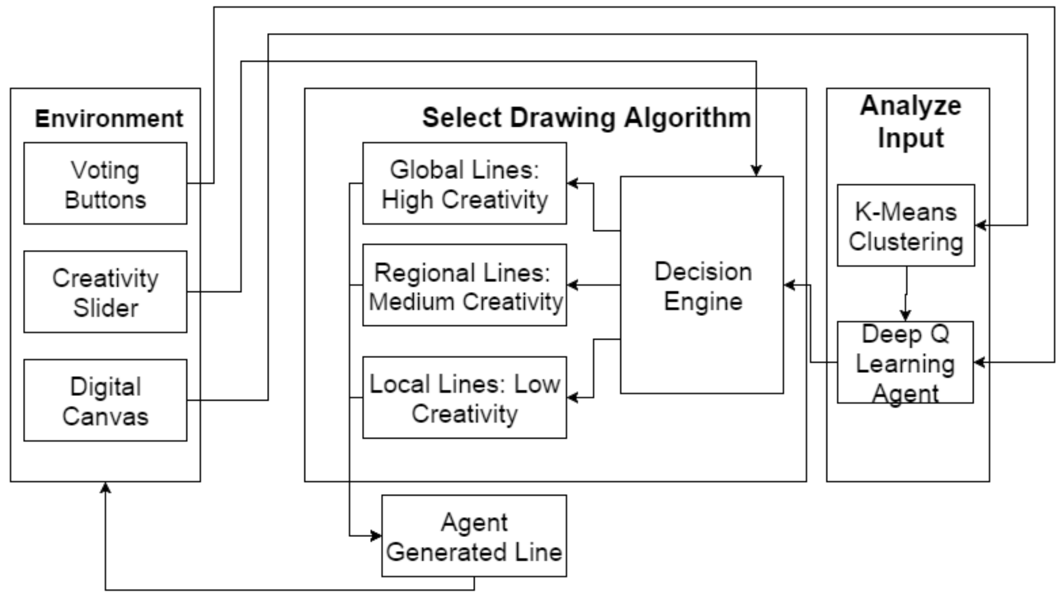


Figure 15: System Architecture for Early System Prototype Utilizing Reactive Algorithms

The creativity slider in the interface of the Drawing Apprentice constrains which algorithms the system will choose from when reacting to the users lines. The creativity level roughly corresponds to the general definition of the term creativity in the creativity literature as: novelty, value, and surprise. At low levels of creativity (slider is between 0-33), the system will produce lines that slightly alter the users input lines without much change. For example, in Figure 16-Algo-2, the system introduces some noise or perturbation to the input line and then redraws it. This type of contribution is almost identical to the users input line and therefore is not very novel or surprising. Other algorithms at the low level of creativity mimic the users line, but slightly translate or offset it, e.g. Algo-3 in Figure 16.

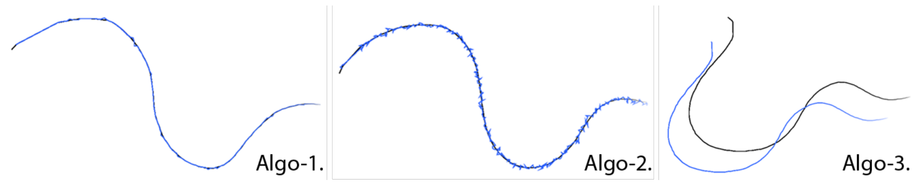


Figure 16: The drawing results from the algorithm 1- 3 for low-creativity level (black lines: human, blue lines: agent

Unique and defining features of input lines set are determined by clustering the data points in the input lines and sending that cluster data into the neural network (see Analyze Input box in Figure 15). This allows the neural network to derive its own classifications scheme based on the data it has been given. The system was seeded with 12 experimental line transformational algorithms, including simple functions, such as translation, scaling, rotation, as well as more complex techniques that change the individual features of the input lines to create new lines that retain a similarity to the input lines. These more complex transformations are achieved by determining a set of equations to describe the input line and then tweaking individual coefficients in the equations to produce a similar line (see Figures 17 and 18).

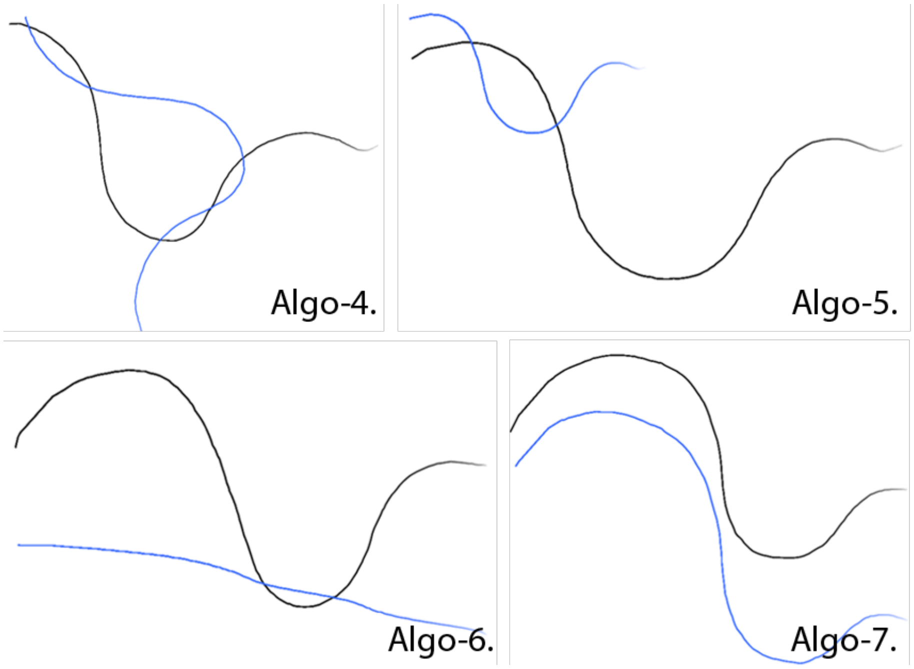


Figure 17: The drawing results from the algorithm 4-7 for mid-creativity level (black lines: human, blue lines: agent)

When the system is set to medium creativity (slider is situated between 33-66), the agent mimics and imitates the users input with transformations, such as rotation (Algo-4 in Figure 17) and scaling (Algo-5 in Figure 17). The agent also mimics the users input line with variations, such as slightly modifying the coefficients of equations that describe the users input line. For example, the height of a curve might change, or the angle at which a corner is made might change. These lines appear near the users input line, but not on top of the line, as most of the algorithms for low creativity are.

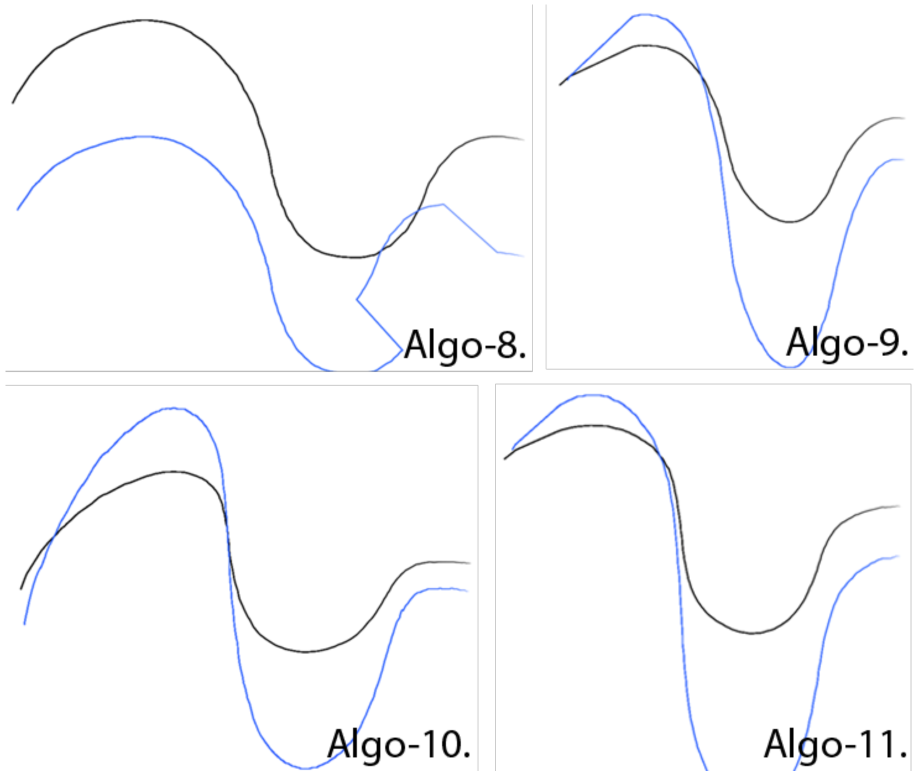


Figure 18: The drawing results from the algorithm 8-11 for high-creativity level (black lines: human, blue lines: agent)

At high levels of creativity (slider is set to 66-100), the system increases the novelty of the lines it produces using a few different techniques. For example, the system employs line mutation algorithms that take the input line and splice a portion of that line with a portion of another line on the canvas to introduce new content into the line (algo-8 of 18). Another algorithm fits a polynomial function to part of the line that can be approximated using a polynomial (i.e. it passes the vertical line test), and then tweaks some coefficients similar to the medium creativity algorithm, but to a more drastic extent, thereby reducing the visual similarity to the input line. Yet another algorithm segments the input line into many different equations (such that it need not pass the vertical line test), and then tweaks the coefficients on some of those line segments while still maintaining the overall shape. For example, if the user drew a square, the system could produce a rectangle.

As shown in Figure 14, the user is provided with up/down voting buttons to give feedback to the system. This feedback informs the system about which algorithms in particular (within each category of creativity, i.e. low, medium, and high) the user prefers. The voting buttons are designed to train the neural network to learn the circumstances under which each type of transformation algorithm should be used. The decision about how to respond to the users input line is a mix of the neural net analysis that compares the current input to previous responses, the creativity value, and the feedback the user has provided on each of the algorithms previously.

4.4.2 Object-Based Drawing Algorithms

In the following sections, we describe the object recognition pipeline, including grouping sketch input, classifying sketch input, selecting what to draw, determining where to draw on the canvas, and finally outputting the contribution on the canvas in an embodied manner (i.e. animating the drawing of the object to simulate human drawing).

The system was initially designed with the intention of collaborating with users in the domain of abstract art. While interacting with the system, many users tend to draw representational objects in their abstract composition. Equipped with only the reactive algorithms, the system reacted to each individual line of an object rather than understanding that this particular group of lines have a relatively well defined relationship and logic that determines the form and placement of lines.

In order to enable the system to effectively interact with users as they engage in representational drawing, we decided to implement an object recognition module where the agent tries to classify what the user is drawing. It classifies the users input lines as a known object and then decides what to draw based on that information shown in Figure 19.

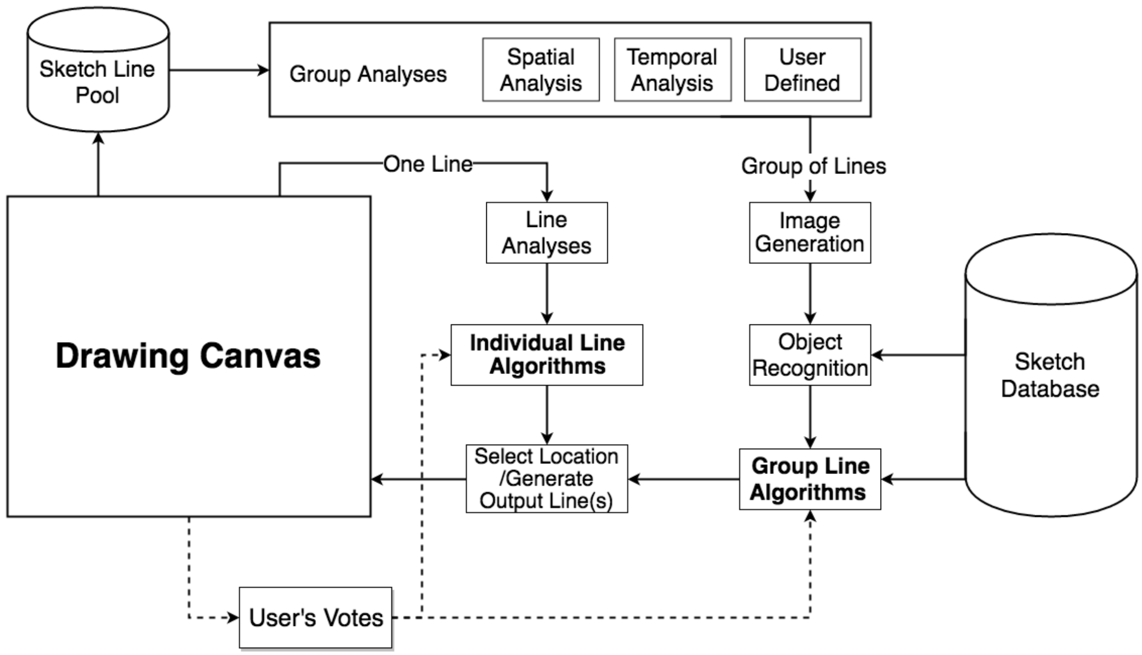


Figure 19: The software architecture for Drawing Apprentice with object recognition and object drawing. The section of the diagram referring to one line contributions was described in earlier sections.

4.4.3 Line Grouping

One of the significant challenges for implementing object recognition in an open-ended drawing application is determining which lines to group together to send to the sketch classification module. One solution would be to offload this task onto the user by having them manually group lines. While this solution is potentially the most accurate, forcing users to manually group every object would significantly disrupt their creative flow, which is an important design consideration for the present application. Automatic and implicit grouping is therefore greatly preferred, but this issue is complex because individuals may begin an object and return to it later. In this case, the input is separated in time but co-located in space, unlike handwriting recognition where one can assume that lines that compose each letter co-occur in both space and time. To address this challenge, we developed a three-pronged solution for grouping sketched lines: (1) time-based implicit grouping, (2) space-based implicit grouping, and (3) explicitly assigned grouping through user input. While the last method needs user input, the first two methods occur without any intervention from users.

In the time-based implicit grouping method, the system starts a timer every time the user lifts their pen from the sketch canvas. If a pre-specified period of time passes between strokes, the system assumes the user has completed a full turn to fully express their idea, and it will mark the last stroke as an end stroke. Based on our observation, we set this interval to 3 seconds. After the time is up, it groups all of the strokes between the previous end stroke and the current end stroke as one turn. These strokes are rendered as a small temporary image isolated from the other strokes on the canvas, and fed into the sketch recognition procedure described in the previous section to classify the sketch.

In the space-based implicit grouping method, the system constructs a quadtree data structure that includes all the points from strokes collected from the human users and AI agent. This quadtree data structure adaptively deepens the depth of the tree structure based on incoming data. Those regions that have a higher density of lines will therefore have a deeper tree structure. Once one particular node is four levels deeper than the average depth of the tree, it returns the area surrounding the node as an area of interest. Then, the system renders an image from the selected strokes in this area similar to the time-based method for sketch recognition and sends this data to the sketch classification module. This approach helps to reduce the computational power required for the common computer vision analyses and helps ensure real-time responses. Between the time- and space-based grouping methods, time-based grouping takes initial precedence, and the space-based grouping performs a secondary check to ensure the new lines are not being added to a previous group based on their spatial proximity to previous groups.

Users can also manually group sketches in the canvas using a lasso tool in the UI to lasso an area of interest for the system to recognize. Similarly, the strokes within the lassoed area will be sent to generate a temporary image, and served as input for sketch recognition. The user can also choose to manually label the object themselves to serve as another ground truth example to help improve the sketch recognition model.

4.4.4 Sketch Classification

We employed convolutional neural networks (CNNs) to classify sketch input due to their recent breakthrough in image and sketch recognition even though methods like Bag-of-Words and SVM worked well in past [51, 187]. CNNs are an ideal candidate to incorporate in the system due to their continued success in recognizing visual information, particularly images and sketches, with a very high accuracy [187]. The structure and functional organization of convolutional neural networks are inspired from the biology of the human eye and vision system. They consist of multiple learnable filters arranged in layers, which each extract relevant features from input images, just as the visual cortex has different layers that each have unique specializations in processing visual information. The cognitive argument for using convolutional neural networks in a co-creative agent is that using such networks would resemble how classification and recognition would occur in the human vision system. For these reasons, we decided to have an end-to-end learning mechanism instead of going the feature engineering route, like Bag-of-Words.

Our sketch classification model was inspired by a VGG neural network due to its recent success in large-scale image recognition. We modified the VGG-CNNs and VGG-19 architecture to suit our task. Since both of these deep neural network models deal with images that contain texture information (encoded using R,G,B channels), we reduced the number of channels to just one as sketches can be represented as binary images [19]. Furthermore, we removed the Local Response Normalization layers from these networks as we found that they work well with images that contain textural information but not well with the task of recognizing sketches [187]. To reduce overfitting, we made use of data augmentation where we randomly flipped horizontally and scaled the training images in addition to using a higher dropout rate of 50

The requirement for having a real time sketch classification engine favored the VGG CNNS model as it has less parameters (and resultantly takes less time to feed-forward) than VGG-19 [155], even though the VGG-19 model provided a greater classification accuracy. Therefore, we decided to utilize the VGG CNN-S architecture. The structure of our VGG models have stacks of convolution layers with smaller filter sizes compared to the Sketch-a-Net architecture with large filter size and high strides [155, 187]. Smaller filter size helps detect local sketching patterns such as crosshatches in conjunction with the overall sketch. The main difference between Sketch-a-Net and our model is that Sketch-a-Net uses a multi-channel and multi-scale pipeline with stroke-ordered training data, whereas our model operates on a single scale and single channel and the strokes in the training data are not ordered. As a result, our current classifier is given a very limited set of information about the sketch. The training of the network was done on the TU-Berlin sketch database, which has 250 categories with each category having 80 different example sketches [51]. During the training phase we split 90 percent of the data into the training set and the remaining 10 percent of the data as test set.

Through our experiments we found that other optimization algorithms such as ADAM, AdaMax and RMSProp did not work that well for training these models [98]. Furthermore, we found through experimentation that the learning rate should be as low as possible, which was 0.001, in order to train the network incrementally. Hence, we made use of Stochastic Gradient Descent with Nesterov Momentum with a learning

Table 1: Comparison of classification accuracy for different state of the art methods employed in non-realtime environments.

Approach	HOG-SVM	Sketch-a-Net	Le-Net	VGG-CNN-S (modified)
Percent Accuracy	56	74.99	55.2	63.9

rate of 0.001 and momentum of 0.9 to train these models. Keeping the learning rate to a minimum helped to counter overfitting to the training data and helped reach an accuracy of 63.95

4.4.5 Object Placement

Determining where the system should draw the target object is a non-trivial task. One of the primary findings of previous user studies pointed to spatial awareness as one of the primary needs for a co-creative drawing agent [34]. The agent should respect the history of the interaction, meaning it should not mess up things that have already been drawn in the past. In this project, we have two main criteria when finding a location for the agent to draw a new object: (1) empty regions where the targeting drawing object would minimally intercept with existing objects; and (2) spatial proximity to the object that was drawn in the last turn. As mentioned previously, all points (both user and agent) are stored in a quadtree data structure for further analyses that is utilized in determining object placement in real-time.

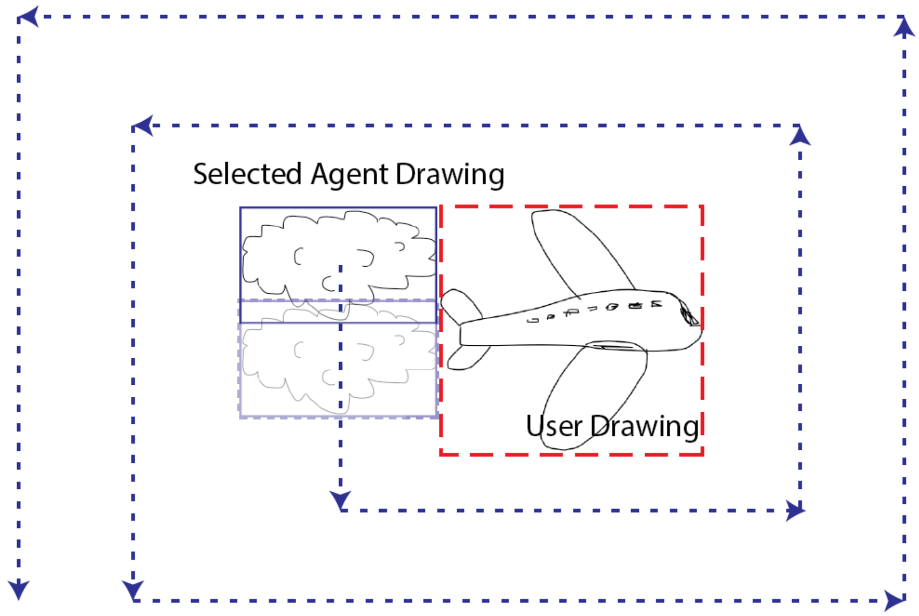


Figure 20: Iterative spatial search procedure to find drawing area near users object.

Once the system detects a turn, it uses the bounding rectangle formed by the users sketches to find an area to draw. As shown in Figure 20, the system iterates through the surrounding locations starting from the top-left corner of the sketch from the users current turn first and then queries the quadtree to get a candidate bounding rectangle containing the least packet points for drawing area. This approach ensures that the target object is drawn as close to the users previous input as possible without drawing on top of existing elements. Figure 21 shows examples of the users sketch and the locations where the system picks for drawing. With the results of turn detection, sketch classification and placement, the system utilizes the following two modes for generating the new sketch objects.

4.4.6 Drawing Similar Objects Mode

In this drawing mode, the system recognizes the users drawn object and then responds with a different representation of that same object, mimicking the users object with intelligent variation. Figure 21-left shows an example where the user drew a chair in the perspective view. The system responded with another chair similar to the original chair. The system uses the t-SNE algorithm [51]) on the visual features extracted by the convolutional neural network to compute the nearest neighbor image in 2-dimensional embedding of the features. Together, the convolutional neural network in conjunction with t-SNE embedding provides the system with the ability to draw visually similar or dissimilar objects (relative to the user input) of the target category. Once the system has determined an object to draw, it selects a location, as described in the object placement section. Once a location has been selected, the system animates the lines in real time to simulate the embodied process of drawing through time.

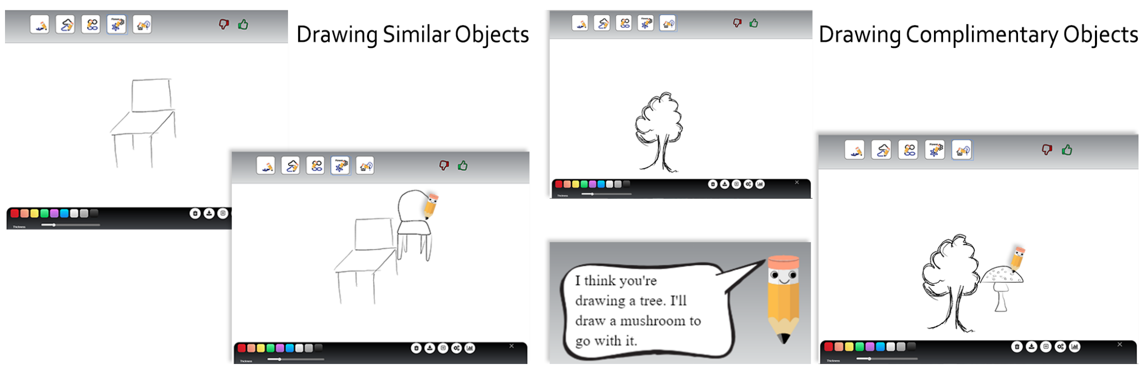


Figure 21: Drawing modes using sketch recognition to draw similar (left) and complimentary (right) objects next to the users most recently drawn object. The agent explicitly expresses what it recognizes and plans to draw (middle).

4.4.7 Drawing Complimentary Objects Mode

In this drawing mode, the system runs through the same object recognition pipeline. However, the procedure diverges when the system determines what object it should draw. Here, the system selects a semantically related category and then picks an object from that category, rather than drawing an object from the same category. The right side of Figure 4 shows that the system recognized a tree has been drawn by the user, then responded with a message in a speech bubble stating its interpretation and planned contribution, and finally drew a mushroom on the canvas. To pick a category, we manually created a dictionary that categorizes the sample sketches into 15 high-level categories (with several sub-categories) based on their semantic meanings. For instance, we group all the animals as one category with marine, bird, and land animals as subcategories.

Ideally, the system should ultimately utilize existing concept nets, such as ConceptNet3 [81] as well as learn new relationships by observing what objects users typically draw together. As shown by the literature on concepts and categories, these elements are dynamic and subject to change based on context and intention [100]. To account for this plasticity, we plan to implement a module that analyzes which objects tend to be drawn together and use this data to inform this algorithm in the future.

4.5 Conclusions

This chapter described the technical implementation of the Drawing Apprentice system. Early versions of the system utilized reactive algorithms that transformed the users input in different ways based on the value of the creativity slider, the type of line, and the users previous stylistic preferences determined through voting. This prototype allowed us to explore the user of machine learning to learn relationships between the users input lines and their preferred responses. Limitations of this approach became apparent through user studies that showed users expected a collaborative partner to understand representational objects they were drawing. To explore integrating sketch understanding into the system, we implemented an object recognition pipeline. We identified challenges for implementing real-time object recognition in the open-ended context of collaborative drawing, and methods for overcoming these challenges. This module was used to develop two new drawing modes that attempt to classify what type of object users are drawing and respond with different versions of that object or semantically related objects.

CHAPTER V

USER STUDY EVALUATION OF DRAWING APPRENTICE

5.1 Summary

This chapter reports on the results of evaluating the Drawing Apprentice system using multiple methods and studies to address the core research questions put forth in this dissertation. These studies include formal laboratory experiments, user study interface evaluations, expert art panels, and informal interactions during public exhibitions of the system. The findings help lay the foundations for understanding the creative needs of users for collaborating with co-creative agents in the open-ended domain of improvisational drawings. We introduce and explain different interaction paradigms for working with co-creative agents that are relevant for unique user groups, such as artists, non-artists, designers, and children. Sense-making strategies that users engage in while interacting with the system are also reported along with a framework that describes when users expectations were violated in ways that did not make sense to the user. Together, these findings help provide an empirically grounded framework describing how co-creative agents can interface with users in an open-ended and improvisational creative context.

5.2 Introduction

Evaluating creativity and creativity support tools is a significant challenge. Many researchers have advocated a mixed-method approach using data triangulation to gain insights on the creative process and technologies interfacing it using a variety of research methods [115]. In the present project, several methods have been employed, including formal laboratory experimentation, user studies, expert art panels, and informal feedback during public demonstrations and art exhibitions. These methods each provide unique data to help understand how users make sense of co-creative agents and what it means to coordinate and collaborate with these systems in the open-ended creative process of drawing.

An iterative testing method was applied throughout the development of the Drawing Apprentice system, meaning multiple versions of the system were evaluated to determine how best to continue development based on user feedback. These innovations include changes in the learning algorithms, drawing algorithms, as well as UI design and functionality. During the course of executing these evaluation, we gained many insights both about the system, how it affects the creative process of users, as well as methods for evaluating co-creative agents in general. Both formative and summative studies were conducted to evaluate the Drawing Apprentice system. The formative studies helped understand how users engage with co-creative agents and the types of creative suppor they expect from a collaborative partner. The summative evaluations more formally compare how collaborating with the system compares to a human wizard controlling the system.

5.2.1 Formative Evaluations

An A-B test was conducted to investigate how UI elements relating to user controls would impact the quality of collaboration. Here, two different interaction paradigms for controlling the agent were presented, namely controlling the agents creativity on a continuous scale versus selecting discrete modes that constrain drawing behavior. The system was also exhibited on multiple occasions to the public during GVU research showcases. These informal interactions provided a naturalistic setting in which to evaluate what users would ultimately want a co-creative system to do, i.e. how they might pragmatically integrate it into their current creative practices. These interactions helped define the creative and collaborative needs for diverse user groups. Additionally, the system was exhibited during an art gallery exhibit called ”Where Technology Meets Art” at a local art gallery called Eyedrum. Informal feedback from this venue was particularly informative since parents and their children engaged the tool and provided extremely valuable feedback about this demographic that prompted a follow-up study focused exclusively on children that is currently ongoing. These interactions revealed significant educational opportunities for the system. Finally, artworks produced with the system, along with process videos of their creation, were submitted to an art competition for evaluation. The videos were evaluated and commented upon by an expert art panel, which helps provide insight into how audiences receive works co-created with an agent, especially highlighting the importance of witnessing the collaboration process as part of the final product.

5.2.2 Summative Evalution

The laboratory experiment compared the drawing behaviors and interaction patterns that emerged when collaborating with the system versus a human in a wizard of oz style study. This study helped to understand the metrics users employ for making sense of the system and what type of interaction dynamics emerge between the current system and a expert wizard controlling the system. The creative sense-making analysis was performed on video data from this study to help quantify the difference between collaboration between the agent and the wizard.

5.3 Formative Evaluations

Co-creative agents require a mixed method evaluation process that can includes both formal laboratory user studies as well as in-situ natural interactions to help understand how users engage the technology in different contexts. We conducted a variety of formative studies to understand the creative needs of different user groups. Since co-creation is such a new computational concept, it was valuable to learn more about how users want to ideally engage co-creative agents, how they interpret the system's actions, and how they think about different components of this system in particular.

These efforts included experimenting with different interface and interaction design approaches to help coordinate open-ended creative collaboration between a human user and co-creative agent during improvisational drawing. In particular, we investigated two different UI paradigms for controlling the agent, a discrete method of selecting drawing modes for the agent versus a continuous scale that controls the agents creativity level.We were interested in how these different UI paradigms would affect the users expectations of the system, their mental model of the agent, and their ability to effectively collaborate with the system. In addition to the two UI paradigms for controlling the agent, we explored concepts for encouraging users to help train the system and understand its reasoning process.

As part of our formative evaluations, we conducted an A-B study asking 10 users to provide reactions and commentary about their experience interacting with both interfaces on an open ended drawing task. Interacting with a creative agent in a collaboration is a novel concept for users. As such, it is critically important to understand what their perception of creativity means with respect to such a co-creative agent. One of the interfaces users were presented with had a slider labeled Agent Creativity. Before the user began interacting with the system, they were asked to explain how they interpreted and understood each element of the interface, including this slider. This provided a means to begin to investigate how users implicitly understand what creativity might mean in the context of a co-creative agent. Participants responses to this query generally fell into three related yet distinct mental models of the systems creativity setting: controlling similarity to user input, user control, and the complexity or intelligence of the agents contributions.

Similarity to their drawn input was by far the most common way that participants conceptualized the agents creativity. Participants generally felt that lower creativity values would correspond to the agent drawing something more similar to their own contribution, and higher values leading to contribution that significantly diverged from their input. For example, P3 commented that higher values of creativity would result in an "abstract version of what I drew." In the same vein, P7 commented that I assume if it is low creativity, it is 100

The next most common conceptualization of creativity referred to two factors: the degree to which the agent understood the users contribution, and the relatively complexity of its responses. At higher creativity settings, participants tended to expect the system to learn what type of object or element the user was drawing. For example, P6 described their idea of high creativity as the system "learns thats [the users input] a person, so if I start drawing a head and neck, and it will start drawing the arm." Here, creativity is correlated to the agents ability to understand and intelligently respond to the users input. Additionally, the complexity of the systems drawing contribution fell into this category, with P6 also guess that "high creativity is drawing three dimensional planes and different things that are a little more advanced." Similarly, P8, after interacting with the system, stated that it "seems the type of drawing is getting more complex" as the creativity was increased.

While it was not as pervasive, two participants thought creativity might refer to how much control they have over the direction of the drawing. When creativity was low, they assumed they would be directing the course of the drawing, whereas if it was high, the agent would have more control over how the drawing progressed. This is exemplified by P8 before he began interacting with the system when he stated "If I put it to the left [zero], I guess he wont be involved at all...all the way right, he will create something." Here, creativity is associated with the perceived autonomy of the agent and its subsequent impact on the development of the artwork.

5.3.1 Modes of Collaboration

During the interaction with the system, participants often made reference to the structure of the collaboration, i.e. the interaction design of the collaboration experience and what type of activity they and the agent would be mutually engaged in. Often, this type of conceptualization relied both on the degree of autonomy of the agent and the type of artistic intention of the user, i.e. whether they had a final end state of the artwork in mind.

Two broad paradigms of collaboration emerged that users employed to make sense of interacting with a co-creative agent: distribution of labor and creative conversation. In more goal oriented contexts, users felt more comfortable providing explicit commands to the agent to accomplish certain tasks. While the user did not want or need precise control over how the agent would accomplish the task, they wanted some control over the constraints of the agents activity. Alternatively, in the creative conversation paradigm, users tended to view the collaboration as a source of inspiration and guide moving forward. A creative conversation de-emphasizes control over the agent in favor of a fun an entertaining creative dialog that is advanced through the mutually influential interaction between user and agent.

5.3.1.1 Distribution of Labor

Within the collaboration framework of distribution of labor, there were several distinct types of activities users wanted the system to accomplish. The most frequently mentioned activity was the application of patterns. In this context, a pattern is the application of some base element repeated throughout a specified space. This space can be bounded (like stripes on an animal) or unbounded (like clouds in the sky). Users wanted the system to engage in pattern both concurrently them as well as on its own, perhaps after demonstrating the base pattern that should be applied more globally to an area.

Key considerations for the applications of patterns include determining the base element, internal constraints for its repetition, and whether the pattern should have a hard or diffused boundary. Internal constraints for pattern application relate to the spatial distribution of base elements, i.e. is the spacing between visual elements uniform (like stripes) or random (like grass). Further, some patterns should vary the base element. For example, with clouds, the base element of a single cloud should be varied in different ways to provide a diverse and interesting skyscape full of clouds in different shapes and sizes.

Users also wanted the system to finish visual elements and objects they started. Partial object completion and partial scene completion are perceived as intuitive modes of co-creation. For example, if the user draws the wheel of a bicycle, they expect the system to finish that object, or at the very least, contribute some elements that lead to the eventual completion of that object. Similarly, if the user drew a bike, the system might draw a semantically related item, such as a hill. Partial scene completion can help the system and user develop narrative cohesion for the artwork, such as a bike that is about to go up a hill. This activity could be concurrent, with both the user and agent contributing during the same time interval (through turns or synchronously). Interestingly, though, users were also interested in beginning an object and then leaving that task for the agent to finish while they begin something new. Users expected the agent to improve upon their initial object formulation as well, furthering the aesthetic appeal of the target object.

5.3.1.2 Interaction Dynamics of Distribution of Labor

These findings indicate that users expect collaborating with a co-creative agent to be less like working with a unique personality and more like providing tasks and constraints for a semi-autonomous agent. This approach suggests effective distribution of labor collaborations might feature multiple agents that can be assigned tasks concurrently while the user works on developing new content for the artwork.

Considering distribution of labor through the lens of participatory sense-making helps understand the nature of the interaction dynamics and meaning construction at play in these scenario. Here, the autonomy of the co-creative agent is reduced and the user is made the de-facto leader of the collaboration by providing boundaries and guidelines within which the agent should operate. Within those bounds, however, the agent still has the creative liberty to decide precisely how the task should be accomplished. The users are not interested in completely removing the autonomy of the agent, but their experience with the current agent prompted them to desire more explicit control over the types of activities of the agent.

Given the reduced autonomy of the agent in this paradigm, this interaction would not be defined as a mutually co-regulated coupling in enactive terms of participatory sense-making. The agent is coupled to the user, but the reverse is not true, meaning that the intention and actions of the user are disproportionately determining the overall trajectory of the collaboration. This type of exchange would theoretically limit the degree of emergent meaning that can potentially be experienced by the user. Instead of a computer colleague, this type of co-creative agent should be considered more as a creative assistant that operates intelligently within bounds determined by the user.

The prevalence and demand for a creative assistant by users communicates two important insights. First, when an agent has full autonomy over its choices and those choices violate the expectations of the user to an unacceptable extent, the user desires more control over the process. The Drawing Apprentice is not yet at the same level of artistic and collaborative competence as a human user, but it was given autonomy to determine how to behave throughout the interaction. As a result, users might desire more control by default. Second, trying to mimic human collaboration through participatory sense-making, while a promising objective to advance computational creativity, may not be the most useful and enjoyable way to interact with a co-creative agent for end users.

5.3.1.3 Creative Conversation

The concept of a creative conversation is more aligned with the original intention for the Drawing Apprentice system. Ideally conceived as a partner that both builds upon as well as challenges emerging meaning structures to prompt creative re-interpretations and emergent meaning. Expected to push individuals out of these creative comfort zones and introduce new visual ideas, the co-creative agent was expected to encourage users to focus more on interacting with the agent than achieving an overall creative product. However, as shown in the previous section, users often have their own conception of what doing art and collaborating should mean, which often involves the concept of distribution of labor. While not ideal in facilitating participatory sense-making and emergent creativity, distribution of labor has the potential to creatively engage individuals in a completely novel way. Users did, however, find significant value by framing their interaction with the agent as a creative conversation. In particular there were three unique activities users described within this idea of creative conversations: creative re-interpretation, one-to-one correspondences, and introducing new ideas.

5.3.1.4 Creative Re-Interpretation

The concept of creative re-interpretation is similar to the idea of a conceptual shift described in the creativity and cognition literature (Nersessian). Here, a set of stimuli that was initially conceptualized as one object or grouping, is viewed through a new lens to see something different from the same or similar input. Viewed as a critical ingredient of solving insight creativity problems, re-interpretation often includes an element of surprise and as the shift is by its nature unexpected.

In collaboration, users would often re-interpret the overall meaning of a scene given the agents last contribution, which would lead to a novel conceptualization. For example, bicycle wheels can become glasses, or a mountain range could become the teeth of a monster. In this scenario, the role of both the user and the agent is to try to see new possibilities in the current input. These moments of re-interpretation can have a positive affective impact on the user in the form of surprise and entertainment. The more unexpected the re-interpretation, the greater the degree of surprise experienced by the user. Of particular interest in this study, though, is that the user was responsible for interpreting what it assumed the agent interpreted the scene as. In reality, the agent was not intentionally re-interpreting the scene, but in fact through its reaction to the user, some of the marks it generated appeared as changing the interpretation of the scene in one way or another.

There are two important considerations when designing to support this type of collaboration. First, the co-creative agent should explicitly re-interpret the scene (or a subset of the scene), meaning the agents actions should not be happenstance and only appear to re-interpret the scene to the user, but in fact, the agent should have some shift in understanding how elements in the scene relate to each other using unique mental models. Second, determining precisely when to re-interpret a scene is a non-trivial task as re-interpreting too often can confuse the user and lead to frustration. We might make a comparison to trying to build something on ever-shifting sands. Some degree of dynamism is required and appreciated during this type of interaction, but there is a certain degree of stability that is required for each individual idea before the re-interpretation can have the desired surprise effect upon the user. One user mentioned that it seemed like the agent was trying to re-interpret the drawing after every turn, which became exhausting trying to constantly make sense of what the agent was doing and then determine how next to transform the scene.

5.3.1.5 One-to-One Correspondences

Another unique way to drive the collaboration as a creative conversation is to employ a type of one-to-one correspondence between the agent and users turns. For example, if the user is working on drawing a house, the agent could begin to draw its own house is a different space. Then, each time the user adds a component to their house, the agent would add a corresponding component to its own house. It need not necessarily be the same component, but some growth from the initial form provides the felt sense that the agent is getting ideas from the user and incorporating those ideas into its own activity.

5.3.1.6 Introducing New Ideas

Instead of completely re-interpreting the entire scene, a co-creative agent can also introduce completely new ideas onto the canvas to help inspire the user. Instead of repealing the unrelated contribution (as might be done with an undo button in the distribution of labor collaboration paradigm), here, the user is challenged to integrate that new element into the existing artwork. For example, maybe the system draws a fork in the sky. Then, the user might expand their definition of what that sky is and draw other utensils raining down from the sky. Without the seemingly random contribution of a fork in the sky, the user would never have reached the conceptualization of utensil rain that arose through making sense of the agents contribution. In this scenario, it is critical that users are in a dynamic and flexible mindset rather than rigidly fixed on a final end state for the artwork. When the interaction is framed as a creative conversation, these types of contributions are interpreted through this type of sense making lense where users strive to create a coherent narrative thread through a seemingly chaotic artistic space.

5.3.1.7 Interaction Dynamics of Creative Conversation

The interaction dynamics of a creative conversation correspond to a co-regulated coupling during which participatory sense-making emerges. Here, the agent and user are granted full autonomy to change the entire meaning and direction of the drawing. As a result, there are many more opportunities for emergent meaning and exploring new creative ideas. In this experience, the end product is not fixed and is typically not as great a concern as being creative and imaginative in the moment.

The primary finding in terms of the interaction design of the system was that the timing of turns was particularly important for coordinating collaboration. There were 10 unique and unelicited comments relating to uncertainty about when, exactly, the agent was going to respond. Though users were informed that the system would respond approximately 4 seconds after their turn was finished, there was no visual indication about where the agent was in that countdown. Further, some users wanted to vary this response time in order to facilitate a more fluid and natural collaboration. For example, one user mentioned the desire to explicitly delineate turns, such as starting and stopping the agents turns manually. Another user commented that the agent has the potential to interrupt the creative flow given the unpredictability of when it would respond to the user.

These findings highlight the importance of transparency and customization in turn-taking parameters governing the collaboration. For example, a simple countdown meter showing how long before the agent will contribute to the canvas would help users orient their expectations and facilitate a more natural creative dialog with the agent.

5.3.2 Practice-Based Evaluations

The Drawing Apprentice system was exhibited during public demonstrations as well as interactive art shows to many people throughout the years of its development, such as GVU Demo Day at Georgia Tech, interactive art exhibits at Eyedrum gallery in Atlanta, and creative technology education events for children. The feedback received during these events was particularly valuable since many different types of users interacted with the system, including artists, non-artists, designers, and children. This section will describe the informal findings from these evaluation activities.

5.3.2.1 Non-Artist Perspective

Non-artists that engaged with the system largely viewed it as a form of art therapy that would encourage them to engage in artistic creativity more often. In particular, some users suggested this type of technology could potentially serve as a more sophisticated version of adult coloring books. In recent years, adult coloring books have gained in popularity. Individuals report that adult coloring books help them relax as well as stimulate creative cognitive processes that may not be utilized in their everyday lives. By building on the users contribution and offering new ideas to explore, the Drawing Apprentice has the potential to engage non-artists in a creative conversation through drawing that is both entertaining and cognitively stimulating. These users also mentioned the benefit of being able to collaborate with their friends (i.e. multiplayer mode in the system) as well as the agent on the canvas, thereby turning the adult coloring book activity into a social process that may further encourage artistic creativity by way of social bonding.

In terms of system functionality, these users seemed to thoroughly enjoy instances where the system responded to their input with semantically related objects, i.e. users draw and eye and the system draws eyeglasses. This type of interaction aligns well with the creative conversation metaphor that has been emerging in our research related to non-artists collaborating with a co-creative agent. The dialogical turn-taking component of this creative conversation prompts users to respond and continue the interaction. This dialog helps prevent task-abandonment, which can be a significant barrier to non-artists that lack confidence in their ability to creatively express themselves. Instead of worrying about the final outcome, users are more focused on responding to their partner with an interesting and creative contribution that builds on what has been previously contributed.

5.3.2.2 Artist Perspective

Artists that engaged with the system provided a perspective that shared some needs with non-artists and also diverged significantly in terms of accomplishing goal-oriented tasks. The needs of artists seem to fall largely into two categories: creative inspiration and creative assistance. Creative inspiration was similar to the idea of a creative conversation described above, with additional creative needs for artists related to removing creative blocks, thinking outside of the box, and helping to explore new artistic ideas and styles. Unlike non-artists, however, artists were strongly concerned with how the system might help them draw better and get things done. Surprisingly, many of the artists that interacted with the system had little confidence in their drawing skills, such as being able to accurately represent objects in an aesthetically pleasing manner. As a result, they wanted the system to help them draw more aesthetically pleasing versions of elements they began to draw.

Artists were interested in enabling the system to accomplish partial object completion, i.e. the user draws a bike tire and the system completes the rest of the bicycle. Going one step further, some artists wanted their initial sketch to be completely replaced by a high-fidelity and aesthetically pleasing version of the object they were trying to draw on the canvas. This type of task-oriented collaboration falls into the category of creative assistance since there is a relatively clear end state that the user envisions and they would like the system to help them reach that state with less time and effort. With respect to getting things done, artists wanted the system to attempt to understand and predict their creative trajectory, where they are headed in terms of the artwork based on what has been done thus far. Importantly, artists want to have a means of viewing and manipulating the creative trajectory the system calculates.

We might use the metaphor of a captain steering a ship to help understand how these artists want the system to interface with their creative process. This metaphor of a creative captain has three components that relate to defining the destination, navigating the course, and directing tasks of the crew. As the captain of a co-creative collaboration, artists want to:

- Define a destination before they embark on their creative journey, meaning select a style or representational objective at the onset, though the goal can change throughout the process.

- Navigate the course by steering the general direction of the collaboration using visualizations of the predicted path the artwork will follow given activities to this point, i.e. plotting the creative trajectory.

- Direct tasks of the crew by monitoring and providing feedback and guidance to the co-creative agent on an as-needed basis, ideally minimizing the amount of attention required, while still enabling fine grained activity tuning.

The captain and ship metaphor is interesting in this context because artists wants to have clear control over the direction of creative process, but they dont necessarily want to micro-manage the agent. In this type of distribution of labor collaboration, the artist wants the agent to work on a type of autopilot that can easily be understood and tweaked.

5.3.2.3 Designer Perspective

Designers that engaged with the system had a completely different perspective than both artists and non-artists related to ideation. Design students described a process they frequently engaged in during the divergent exploration phase of their design process called pair brainstorming. During this process, two individuals engage in a collaborative design session where they each come up with different versions of a target design. For example, if students were instructed to design a chair, the pair would engage in a dialogical process whereby they each came up with different versions of the chair or different perspectives and scales of the chair. This type of brainstorming helps designers to fully explore the design space and help understand the design problem. They noted that the Drawing Apprentice system could perform the role of their partner so they may engage in this productive form of collaborative brainstorming more often and without having to find a partner to engage in the task. In particular, these designers liked how the system would mimic their designs with slight alterations in unexpected ways. Additionally, the designers found it interesting when the system recognized the object they were trying to draw (e.g. a chair) and drew a completely different type of chair. This type of interaction could help the users fully explore the design space without requiring another individual to stimulate their creativity with unexpected contributions.

5.3.2.4 Child Perspective

When children attempt to draw an object, they may pause for significant periods of time while they think about how to represent their idea. The current system design begins to react to the users input after about 4 seconds have since the users last stroke. In many cases, the child was not done drawing their intended contribution, yet the agent started drawing anyway. As a result, children require either longer turns or a new model for controlling when turns occur.

The system is currently designed in a way that largely follows the user, meaning that the systems turn begins after the user completes their turn. With young children, it seems the reverse might also have an interesting effect on their creative process.

Children were observed mimicking elements introduced by the system. For example, when the system drew a car, this inspired the child to add circles onto a box he had created earlier. This suggests that young children may benefit from a guided drawing experience that intentionally introduces new visual ideas and concepts into the drawing in a way that is meant to generate new ideas for the child. This guided creativity also ties into the idea of theming the session with a certain genre or narrative thread, such as drawing a forest or drawing a city.

Children seem to be more engaged in the process when there is a clear direction and creative goal, such as creating a specific scene that has been agreed upon at the outset of the drawing. During the demonstration, the parents were observed discussing this type of artistic goal with their child to help provide a creative prompt for subsequent items. The parent was also responsible for making sense of the agents contribution within some consistent semantic framework. In this sense, the parent provided a story to help contextualize the contributions of both the child and the system. Enabling the system to utilize this type of reasoning in the generation and description of actions may help sustain creative engagement for children as well as provide a clear path forward for their continued involvement.

Introducing a direct means of communication about the systems reasoning process provided critical feedback for users that significantly increased engagement with the system. However, the current implementation uses a speech bubble with text that explains the reasoning process of the agent, in terms of what it thought the user was drawing as well as what it plans to draw. This speech bubble seemed to increase the affective bond between the user and system through personification and transparency in the reasoning process. For young children that cannot yet read, spoken language paired with text output may be more appropriate.

One interesting practice that emerged was for parents to describe what type of object the system was drawing to their child as well as providing narrative details about that object in the context of the scene. For example, if the child drew a house and the system responded with a dog on the opposite side of the canvas, a parent might comment that the dog is trying to find its way home. This type of response contains several interesting educational opportunities. First, labeling a drawn object may help young children increase their visual lexicon by being able to either learn a new object and its visual representation, similar to educational books that provide pictures as well as labels for the object (e.g. A is for apple). Additionally, these object-label pairings might help to reinforce previously known concepts, such as dog, by presenting a new representational instance of that category of objects. The parents statement also encodes some semantic relations about dogs, i.e. dogs have homes, and they can try to find their way back to their homes, especially when the distance between the dog and home is great. Including some narrative context can increase creative engagement as well as provide a compelling educational opportunity for young children.

5.3.2.5 Providing Feedback to Co-Creative Agent

There were some general remarks across these user groups that related to voting system and way of providing feedback to the system. In general, users wanted the voting and user feedback system to be tied to their implicit drawing behaviors and actions rather than having to provide explicit feedback in the form of binary voting. Many users mentioned manipulating the agents contribution in various ways in a way that could provide feedback to help train the machine learning component of the system. For example, users wanted the ability to scale, move, and remove the agents contributions. Each of these actions could serve to inform the system about its contribution.

5.3.3 Expert Panel Evaluation

As part of the practice-based evaluation of the Drawing Apprentice system, the system was employed to produce artworks through collaboration. This type of evaluation helps to understand what the system is capable of given a user that is intimately familiar with its functionality. In the present evaluation, the lead researcher for the project used the tool for approximately one hour to complete a detailed abstract drawing with the system. That drawing was then submitted to the Clough Student Art Competition at Georgia Tech in the digital art category. Submissions to this category included all digital mediums such as Photoshop. Thus, the artwork was in competition with purely human generated art as well as other procedurally generated artworks. Along with the final product, a video of the collaboration process was submitted as supplementary material. The final product submitted can be seen in Figure 22.

Figure 22: Collaborative artwork donw with the Drawing Apprentice that won the Digital Art category at the Clough Student Art Competition and Georgia Tech in 2015.

Figure 22 shows four time lapse images depicting the process of creation starting from the top left and ending in the lower right. In the three process images, the systems lines are shown in blue and the users lines are shown in black. In the final artwork, both the agent and users lines are shown in black, which is how the user perceived the drawing throughout the process.

The art submission won the digital art category of the art competition, thus demonstrating the value of such a system for generating genuinely aesthetically pleasing products. Though the system was designed with the intention of encouraging collaboration and shared meaning construction, it is important to note its potential for creating polished creative products in the hands of an experienced artist.

The expert panel provided an evaluation of the artwork as follows: The act of collaboration between a person and the enactive agent (presented live or through video documentation) is a visually exciting back-and-forth exchange worth watching, or perhaps even participating in yourself. Stepping Stones excellently demonstrates the way technology can partner in an artists creative process, as well transform passive viewers into collaborators themselves [7]. The dialectic process of improvising and collaborating with a co-creative agent itself is a product when viewed through time. The back and forth exchange engaged viewers in a narrative manner, watching the progress unfold and anticipating what might come next.

5.4 Summative Evaluation

We hypothesize that a system that mimics and builds on the contributions of users in a real-time drawing collaboration can impact the sense-making process of art creation with similar benefits as participatory sense-making exhibited in open-ended creative collaboration between two people. This user study was designed to help understand participatory sense-making in the domain of collaborative drawing and delineate critical mechanisms that foster it. We investigate to what extent users can effectively work with the Drawing Apprentice in a way that enables the user to interactively and co-creatively build artistic meaning as the artwork develops.

5.4.1 Study Design

For this study, we had 7 participants, 4 female, and 3 male with an average age of 25 (ranging from 20-45) recruited from the student population at Georgia Tech. The artistic experience of the group was generally categorized as novice, with an average of 2.15 on a 5 point scale ranging from no artistic experience to 5 years of professional practice in the field. The data generated from the study included video recordings (see Figure 6), the transcribed audio data from the retrospective protocol analysis, the log data from the system, and the survey data.

The experiment was divided into two phases that each included a 12-minute collaborative drawing task, a retrospective protocol analysis, and a survey about the participants experience interacting with the system. A non-collaborative drawing task was not included because drawing independently lacks any social coordination and participatory sense-making, which were the focus of the current study. Each experimental session lasted approximately one hour. The experiments were conducted using a Microsoft Surface tablet and a capacitive pen as input to the device. The Drawing Apprentice system was running as a web application and expanded to full screen.

Participants were first oriented with the basic drawing features of the interface (line thickness, color selection, input method), as well as the unique features of the Drawing Apprentice system, such as the voting buttons and creativity slider. The experimenter described how each vote helped the system understand what the user liked, and the creativity slider controlled how creative the agent was, with 0 being less creative and 100 being the most creative. Participants were then given an open-ended prompt to collaborate with the system for 12 minutes to create a drawing. One drawing task was collaborating with the Drawing Apprentice system (referred to as the agent condition), while the other task was collaborating with a Wizard of Oz agent being controlled by an expert human artist (referred to as the WoZ condition). The interface was the same in both conditions, and participants were not aware of which condition they were experiencing. They also were unaware that an expert human was controlling the system in the WoZ condition. The experimental conditions were randomly ordered to account for learning effects.

In the WoZ condition, an expert human artist controlled the systems drawing contributions from another room. The expert artist was one of the researchers from the team that has been collaborating with novices and experts in the domain of abstract drawing for over ten years. The expert used approximately the same interaction dynamics as the system, waiting 2 seconds after the user was finished drawing to begin the turn, as well as drawing approximately the same amount of lines as the user's last turn. The expert's responses were calibrated using the creativity slider, with low creativity turns closely following the participants lines, while high creativity resulted in more novel contributions.

During the retrospective protocol analysis, the experimenter prompted the user to explain their thought process throughout the video walkthrough. The survey focused on evaluating the quality of the collaboration and how influential their collaboration was in defining artistic goals. This section presents the results from the retrospective protocol analysis. We performed a qualitative data analysis on the transcribed interview data using thematic analysis targeting language relating to key themes about participatory sense-making.

A key factor in this analysis is determining why contributions seem to make sense to the user or not make sense to the user. We expect contributions that appear completely erratic to be too far outside the meaning structures developed by the user to be integrated, while contributions that are on the fringe of the users current meaning have the potential to expand visual ideas and guide interactions in-the-moment. Further, the manner in which interaction occurs can impact whether the agent appears to coordinate with the user, i.e. the rhythm of turn taking, the speed of lines drawn, the size of lines, etc.

Supporting evidence for participatory sense-making comes from participants reporting that the agent was able to contribute to the drawing in the following meaningful ways: (1) build on the contributions of the user, (2) demonstrate a certain degree of coordination and mutual sense-making with the user. In particular, we are interested in whether the agent can help shape these meaning structures such that creatively engaging, unexpected, and surprising meaning emerges through interactions over time.

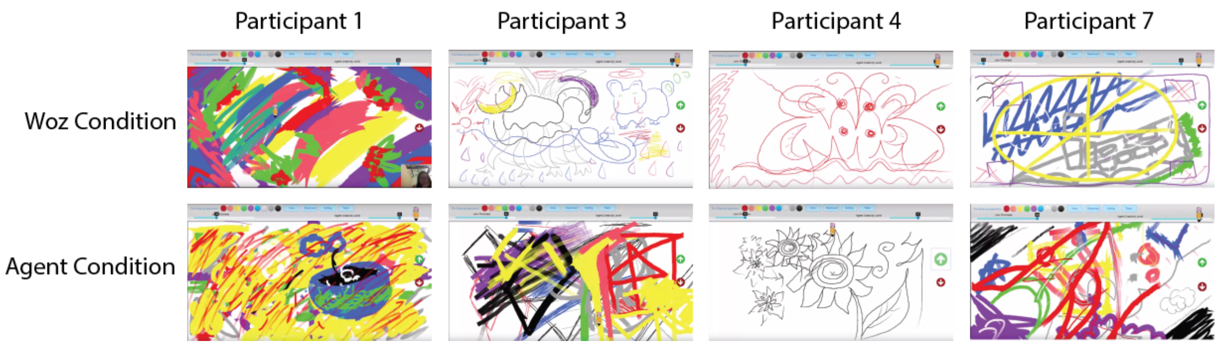


Figure 23: Example artworks created during Drawing Apprentice user study.

5.4.2 Wizard of Oz Condition

The first step in the analysis was to delineate how participants described participatory sense-making in the case of human collaboration (WoZ experimental condition). In this circumstance, the thematic analysis focused on concepts related to a structural coupling between the collaborators, during which there was a mutual co-regulation of the activity, meaning both the human and the system defined basic units of meaning that were extended and added on by each participant in a way that would not have occurred without the collaboration.

5.4.2.1 Making Sense of the Agent

P2 described how the agent influenced the creative process, reporting "I'm definitely taking into account what the agent is doing and some combination of trying to figure out how to sort of control the agent, and also work with it, because it did things that I wouldnt necessarily expect, which is cool." P5 describes how echoing and mirroring actions helped anticipate and make sense of the agent, stating:

> "The system tended to echo what I did, and mirror the patterns I

> made, sometimes in different locations on the screen, but I was able to pretty quickly anticipate the kind of move, not exactly what it would do, but to the point where I was pleased with some of them, or displeased..."

P5 elaborated this sentiment by describing he was able to find a comfort zone in which the agent would make predictable contributions, "At that point, I liked what I expected it to do. That didnt really surprise me there, but in a comforting kind of way." While unexpected contributions may have inspired the user, achieving basic structural couplings enabled a comfort zone to emerge that provided some stability and predictability in the collaboration that helped contribute to the feeling of a dialog with the system.

5.4.2.2 Interaction Dynamics

P3 described the experience of WoZ collaboration in distinct categories that help shed light on how the interaction was conceptualized:

> "There was like me teaching it, or it copies me, or he comes up with something cool, and I copy it, or we were like collaborating on something we both know whats going on. I guess there is just another thing that is random, like I dont know what to draw, and Ill do something random, or he doesnt know what to follow and he does something random."

As P3 worked to make sense of the agents behavior, they noted distinct modes of interaction that included both the agent and user copying, working on a joint activity, as well as injecting novel or random contributions to help move the collaboration forward while the user or system were unsure how to contribute. Examples of joint activities include coloring (i.e. filling a shape with one color) as well as one of the collaborators (agent or user) completing each others thoughts. Achieving joint activity relied heavily on mutual spatial awareness, which is detailed later in the Sense-Making Evaluation section.

The notion of participants copying the agent was introduced again with a strongly positive association from P4. There was a sentiment that when the participant chooses to copy the agent, it signified the agents contribution were accepted and integrated into the users current artistic intention, i.e. structural coupling. P4 exemplied this sentiment in the following passage:

> "It was really cool. It is taking the accent points from the lines and making them more accented. So I was like OK, Im going to copy you.I think throughout this interaction, there are times when the system lead me instead of the other way around... I felt a little like the system. I was like: am I just copying it?"

The system and user were taking turns leading the interaction and suggesting new content and activities in which to engage, demonstrating a mutual co-regulation in the participatory sense-making process. Once users made sense of how the system interacts, the users engaged in a playful process of exploring the boundaries and challenging the system with increasingly difficult inputs. For example, P4 challenged the agent, stating "So, I was like OH! It looks like a butterfly, and I was like haha system! What are you going to do now? So I draw in moons, and I was like OK, what are you going to do?" Instead of worrying about the outcome of the artwork, P4 was creatively engaged by the dialog that emerged with the system. This interaction inspires the participant to generate lines that test the limits of the system.

5.4.2.3 Emergent Meaning

P1 described the process of discovering emergent meaning as a result of working with the agent, saying

> "I drew that, and it started out as random stuff, and it reminded me of a flower or a star, and then the agent drew that, and it reminded me of a halbird, so I extended it, and then the agent extended it further."

In this instance, the collaborators formed a brief structural coupling during which the user and agent were both contributing ideas and building onto a core idea that emerged through interaction. P6 also reports meaning emerging through a process of co-regulated structural coupling when he describes redefining his goals about the agents contributions,

> ”I wanted to make a wagon, a cart thing, but then it made something like this. Then, I made it a bed instead. Then, it did this really nice hatching on the pillow, which I liked.”

Here, P6 originally intended to draw a cart. However, based on the systems contribution, he creatively reinterpreted that goal as a bed, which was further elaborated by the system.

5.4.3 Agent Collaboration Condition

Next, we describe some illustrative examples during which participants reported participatory sense-making during collaboration with the co-creative agent. Then, we will list some of the most critical evaluation metrics users reported for determining whether or not the agents contributions make sense.

5.4.3.1 Making Sense of the Agent

P4 described how she developed a strategy for anticipating the agents response through experimentation,

> ”I decided to make a spiral, and see if the AI would continue my spiral, but then it made it elsewhere, so I was like thats cool...so I drew a different spiral, to see what it would do, and they did the same thing.so I was like maybe I can use that and replicate smaller things, so I would make a flower, so I anticipated what it did.”

This quote shows evidence that P4 began to make sense of the agents reactions. P5 also discovered types of actions that elicited good responses, saying "I did a lot of these Bezier curves because I liked the way it tried to answer me with those...." This type of predictability was critical for the sense-making process in artistic collaboration.

5.4.3.2 Interaction Dynamics

Users were able to couple their behavior with the system, but as P5 describes, this required them to submit to how the system worked in a one-sided manner,

> "This is when I sort of discovered how to work with it on that pattern over there...Im sort of doing what its doing, and were feeding off of each other...I quit fighting it and started collaborating...Im playing more by its rules. Im sort of anticipating what I thought it would do."

This quote demonstrates participatory sense-making, but the process is being regulated mostly by the agent because the participant has to work from the agents contribution in order to achieve collaboration. This is contrasted to the mutually co-regulated process during which each partner takes turns leading and defining new goals, as seen in human collaboration. P5 surrendered a certain degree of autonomy when he decided to play by the agents rules. Some participants were comfortable surrendering this autonomy. However, in some cases, such as P7, dealing with the agent was frustrating, since there was "a lot of prodding him into doing things. It was more about prodding him, and being concerned about him rather than me doing something creative." This type of interaction dynamic could be characterized as trying to lead or control the agent as opposed to engaging in a mutual co-regulation where each party exhibits autonomy. While prodding the agent may reduce the artistic autonomy of users, the dialogical interaction still had the capacity to propel the interaction forward, as P5 describes:

> "I think if you had just given me a blank canvas and told me to draw anything, as a non-artist, I would have quickly given up, because I didnt know what I was doing. The fact that there was a back and forth, and new things were emerging on the screen, made me want to try to answer it, or to try to prod it to coming up with another addition to the drawing. I was probably more engaged than I would be if I was just drawing on a blank canvas by myself."

5.4.3.3 Emergent Meaning

P1 summarized the effect that collaboration with the co-creative agent had on their creative process, saying "It made it less structured. I did more doodling to see how I could incorporate what the pencil was doing and see how I could interact with it." The doodling that P1 described leads to emergent and unpredictable visual elements in the drawing. This type of interaction dynamic helped users discover novel visual ideas. P6 describes how working with the co-creative agent challenged them to move beyond their standard drawing,

> "It felt nice when I had to change my standard drawing, because that's something I always draw. This was more of a challenge for me to say: let's see what I can come up with to draw."

Adapting to and incorporating the agents contributions helped participants draw in ways they would not have done without the collaboration. The challenge for the Drawing Apprentice is determining how to make better unexpected choices that encourage users to incorporate the agents contributions rather than ignore them. Utilizing user feedback is one method to help facilitate this type of coordination.

Table 2: Comparing participatory sense-making between the two collaboration conditions

	Making Sense of Agent	Interaction Dynamics	Emergent Meaning
WoZ Collaboration	User anticipated and predicted agent responses	Co-regulated coupling where leader varies between agent and user	Visual ideas emerge from both user and agent and transform over time
Agent Collaboration	User anticipated and predicted agent responses	User leads and controls agent	User incorporates and adapts to agent responses

5.4.4 Voting Buttons and Creativity Slider

Participants reported being able to make sense of the agent and collaborate with it in both conditions, but the voting buttons and creativity slider were not fully leveraged by most participants. The down vote mechanism was used to discourage the system from behaving in a particular manner, such as the style of its lines. For example, P2 reports:

> "I definitely up voted a lot in the first time, and I down voted a lot more this time. I wanted to try to get rid of the fast lines I couldnt see, which kind of happened...I tried to discourage when it did really really shaky stuff. It kind of helped."

Other participants reported being uncertain about how voting affected the system. For example, when directly asked how voting affected the agent, P1 stated, "I dont know. I dont know if they really affected it." Other participants experimented to a further degree, but still described uncertainty when voting on the agents contributions. P4 reported: "Im not sure if the voting buttons did anything, because I down voted it, but it continued to do what it did." In general, participants expected immediate feedback from a down vote, such as the system immediately stopping its current drawing task. P2 elaborated one reason why voting was confusing, stating:

> "It definitely takes several iterations of a down vote for it to figure out exactly what you are trying to discourage, like the line placement, or what type of thing you are trying to discourage, because there [were] probably several things it considers when trying to place a line."

This comment highlights the fact that each vote could relate to several parameters of the line, such as its placement, the style (shaky, smooth, curvy, etc.), and the shape. Participants were uncertain how their vote would affect the system due to this ambiguity. For example, P3 stated that each system input "has a lot of attributes, I dont know if it discards or has a decreased value of all of those, or I just dont like this particular thing." Providing users with explicit visual feedback about how their vote affected the algorithms would help address this shortcoming of the system. In addition to visual feedback, increased granularity of the voting mechanism was also mentioned by P4:

> "Also, another thought on voting, I guess the binary voting is a little more extreme, I guess [if] there is a scale, not as fine of a resolution of 1-100, but something that gave me 3-5 regions, or even if I wasn't aware of it, and I was hitting on a linear line of good to bad and I could just tap in there, and I could just tap there, because a lot of things were nuanced. It wasn't that this is amazing or this is horrible. I think three would have been too little, too, because a lot of stuff was not neutral either. I like 4 [categories] and neutral is implied as I didnt vote, but still, that would have been distracting to me, but maybe I would have liked it better if it wasnt on the screen, if there were like buttons I could have done with my hands while I watched the screen."

Other participants also described voting as distracting, or similarly described how they forgot about voting when they were deeply immersed in a drawing task. For example, P5 described his experience with voting, saying "I tried to use those [voting buttons] this time to reinforce things I liked, but I found those more distracting." When users were deeply engaged in collaboration, they reported thinking less about the voting and creativity mechanisms. For example, P4 stated "I totally forgot about the creativity slider, and the voting buttons for a while because I was so intrigued by what the system was doing on its own at that level of creativity." These mixed sentiments about the feedback mechanisms provide insights for updating the implementation of voting and feedback that will be outlined in the Design Recommendations section.

5.4.5 Turn Taking and Voting Behavior

Users adopted different types of interaction strategies between their collaboration conditions (agent vs. wizard collaboration) and across the different participants. However, on average, the number of turns between the two conditions is similar across the participants (though there is a significantly higher standard deviation in the agent condition). Tables 3 and 4 summarize the number of turns, average number of lines per turn, voting behaior (e.g. number of upvotes and downvotes), and the summed total of the votes (taking downvotes as a negative value and upvotes as a positive value) in each condition.

Comiparing the the summed votes (summed value of the positive and negative voting) between the agent and wizard collaboration conditions shows that users general upvoted more on average in the wizard collaboratino than the agent collaboration. This result was expected given that the wizard can understand what the user is drawing and respond based on that understanding, while the agent is responding to indivudal lines and the user's preference for different types of line transformations over time.

One interesting correlation from this analysis that is not immediately obvious is

Table 3: Turn taking and voting behavior of participants in the agent collaboration condition of the Drawing Apprentice creativity study

ID	Condition	Num. Turns	Avg. User Lines	Avg. Agent Lines	Up Votes	Down Votes	Total Votes	Vote Sum
P1 S1	Agent	16	8.6	8.2	0	2	2	-2
P2 S2	Agent	45	1	.97	13	8	21	5
P3 S2	Agent	15	5.1	5.2	5	3	8	2
P4 S1	Agent	20	4.4	3.2	5	6	11	-1
P5 S2	Agent	45	2	1.5	15	6	21	9
P6 S1	Agent	61	5.3	4.9	5	4	9	1
P7 S2	Agent	33	1.5	1.5	2	3	5	-1
Avg.		33.5	4	3.6	6.4	4.6	11	1.6
St.dev.		17.6	2.7	2.6	5.5	2.1	7.4	4

Table 4: Turn taking and voting behavior of participants in the wizard of oz collaboration condition of the Drawing Apprentice creativity study

ID	Condition	Num. Turns	Avg. User Lines	Avg. Agent Lines	Up Votes	Down Votes	Total Votes	Vote Sum
P1 S2	Wizard	28	10.6	4.7	0	0	0	0
P2 S1	Wizard	39	1	1.1	12	1	13	11
P3 S1	Wizard	30	3.4	2.3	12	5	17	7
P4 S2	Wizard	31	2.7	2.8	4	0	4	4
P5 S1	Wizard	44	1	1	13	7	20	6
P6 S2	Wizard	42	2.5	1.7	14	0	14	14
P7 S1	Wizard	22	2	2	6	6	12	0
Avg.		33.7	3.3	2.2	8.7	3.8	11.4	6
St.dev.		8.1	3.3	1.2	5	2.8	6.5	5.3

the relationship between the number of lines per turn and voting behavior. There is a weak trend showing that fewer lines per turn resulted in more overall upvotes, as shown in Figure 24.

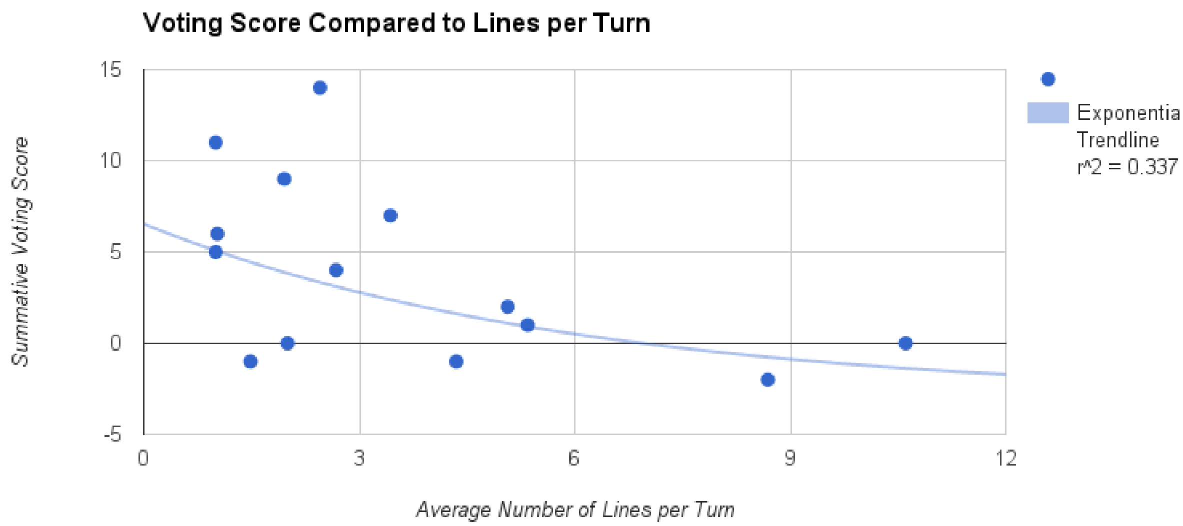


Figure 24: Summed voting score of participants versus the number of lines per turn in both collaboration conditions

There is also a more general correlation between the overall voting behavior and the number of lines per turn, showing an increase in overall votes (whether positive or negative) when there are fewer turns, as shown in Figure 25.

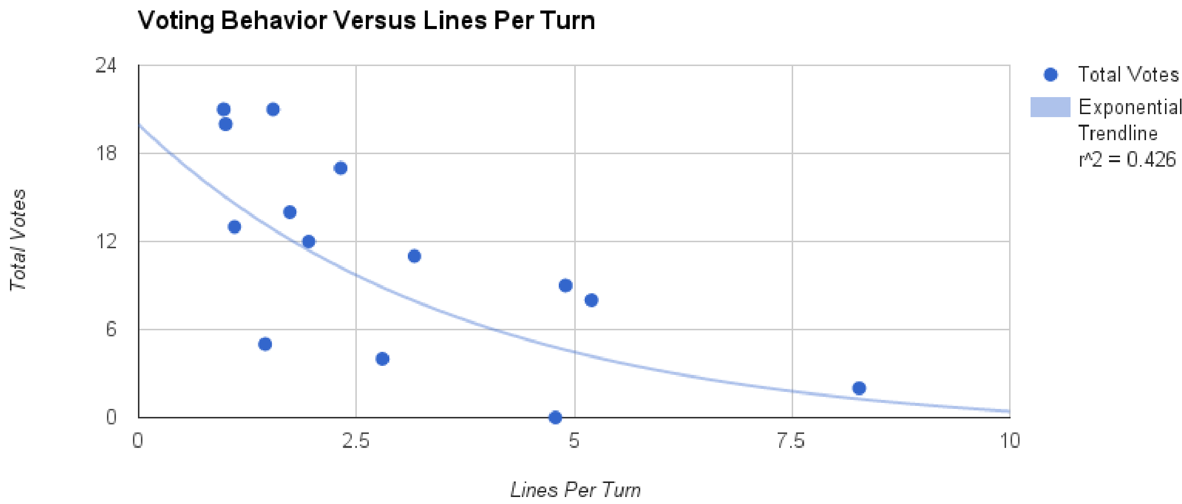


Figure 25: Overall number of votes of users in both conditions compared to the number of lines per turn

One explanation for the correlation between voting score and average lines per turn is that users have lower expectations for the system when they use less lines per turn. Turns with less lines might not express a complete idea for the user, and as a result, the user may be more leniant in interpreting the system's response in that scenario. For example, in the extreme case of P2, the participant only drew one line per turn. In this case, the user and system each created one line per turn and worked together to build up a drawing through these interactions. This participant maintained this turn taking strategy in both conditions, and her voting behavior shows a clear preference for the wizard scenario based on the vote sum (11 for the wizard and 5 for the agent).

P2 created an abstract drawing with the agent that gradually evolved over time in each condition. Other participants drew specific objects during their turns, which often consisted of a higher line count per turn as the participant attempted to express a more complete idea. However, since this version of the system did not have object recogniion, the system was not as well equipeed to respond appropriately to turns with many lines and concrete objects.

5.4.6 Sense-Making Evaluation Metrics

Achieving human-level collaboration is the ultimate goal for the type of co-creative system we designed. To do so, it is critical to understand the nature of participatory sense-making in the domain of artistic collaboration and the metrics users employ to evaluate their partner. Next, we describe key concepts that emerged in the thematic analysis when we focused on contributions that appeared random or erratic in a way that was not beneficial and did not lead to emergent meaning. This dimension of the analysis will help inform the continued development of our system as well as other co-creative agents in the artistic domain. The concepts of spatial awareness, visual similarity, and perceptual logic emerged as particularly important when participants described how they evaluated whether contributions from the system made sense.

5.4.6.1 Spatial Awareness

One of the most common metrics users employed to evaluate whether the systems contributions made sense related to whether the agent exhibited spatial awareness of things that were previously drawn. According to participants, the agent should maintain an awareness of visual elements that have been created previously to not mess up what the users have drawn. Participants became frustrated when they perceived the agent as messing up their contribution rather than building off of it. For example, P1 states: "Usually if I draw actual things, they keep within lines, and you know, you organize the space, and when you have other things encroaching on it, its hard." P4 reported a similar occurrence while collaborating with the agent, stating: "The system would draw on top of other things we drew, when I specifically avoided those thingsso that was annoying." It was obvious that the system did not understand what the users intention was when it drew on top of the participants lines.

Users can readily make sense of the agents contributions when it reacts directly to their last turn. However, users become frustrated when those contributions do not take into consideration what has been done thus far on the canvas. The canvas represents the history of interaction and this plays an especially important role in art making since the artwork is constantly growing through time. An agent that behaves in a purely reactive way will cause frustration by drawing on top of things that have been accomplished by the user previously. Thus, a reactive agent can be made more robust by avoiding drawing on top of previously drawn elements on the page. However, this becomes increasingly difficult as a drawing progresses given that the canvas gradually fills up.

One simple strategy can involve drawing in the blank spaces on the canvas. Yet, as the canvas begins to fill, this strategy becomes less useful. Another strategy could be to constrain the line response by augmenting it with the type of logic of the regions it will overlap with. For example, if the agent has decided to draw a large wavy line that stretches across the entire canvas, it should consider the regions that the line will pass through and work to satify the constraints of each individual region. This introduces another idea of localized constraints, or perceptual logic, that dominate the artwork. Perceptual logic refers to a type of evolving artistic affordances that determine what type of contributions are possible in a given region. It refers to a bounded space of potentialities, i.e. a conceptual search space, that have constraints unique to the arrangement of lines in that particular region.

5.4.6.2 Visual Similarity

The agent's contribution should retain some visual similarity of the users contribution so that the user can help understand the relationship between their contribution and the system's. When the user was able to see the visual similarity between the agent's and their behaviors, they were able to reason through why it made a contribution, which can reduce their frustration if it *messes up* something they drew. For example, P2 states: "This is cool. I like this pattern here. I did that, and he did a more acute angle." The agent's contributions that were on top of the users lines were more acceptable if the user could justify why the agent had created those lines. Typically, visual similarity and close spatial proximity to the users input lines helped provide clues to participants about why the agent drew a particular object. For example, P3 states "I was trying to draw two birds, but he just tries to copy...I didn't really like it, because it's not really a bird anymore, but I get why he did that." It is important to understand the context in which a contribution is accepted even though the participants report not liking it. This acceptance can lead to emergent meaning as users work to continue to transform the content rather than completely abandoning or ignoring the contribution, as is sometimes the case when the user perceives contributions to be too erratic.

5.4.6.3 Perceptual Logic

The agent's lines are evaluated as messing up the participants drawing when the agent drew over top of a previous structure without taking into account any of the features of that structure. When the agent did make contributions within or on top of a previously defined shape, it must adhere to what could be termed the prevailing perceptual logic of that region [8,9]. Perceptual logic is a concept that describes how each region or idea in a drawing has its internal set of rules and mechanism that serve to constrain what type of visual contributions seem relevant and logical for the target region. It could be conceptualized as the local style of a particular region. In some cases, perceptual logic appeared to be based purely on visual relationships, like the repetition of similar patterns (as noted in the previous section on visual similarity), while in other cases, it was based on a semantic definition of the object, such as a house that has different components that are right or wrong. P3 described a perceptual logic of orderly geometric shapes, reporting:

> "I didnt like too much what was going on there because I was trying to actually do like geometric shapes and stuff, so I was thinking something more orderly, so I just moved back and started drawing regular shapes."

Contributions that did not adhere to the perceptual logic of orderly geometric shapes did not make sense to the participant. Another example of perceptual logic is when P5 reported:

> "I didnt like the way it looked. It messed up my swirl, and it wasnt symmetric....if it was going to do it, it should be centered in the spiral, then I might have thought it would have been ok...it seemed arbitrary."

The negative sentiment this participant had for that particular violation shows that sometimes perceptual logic is tied up with an emotional response, evidenced when P5 reiterates his sentiments on the spiral, saying "I really hated that thing inside the spiral, thats what I hated the most."

5.4.7 Sense-Making Analysis

When the agent is reacting to the users last contribution, the autonomy of the agent is low because the human's input turn dictates the amount of lines that will be drawn as well as the timing. In this case, the agent does exhibit autonomy in deciding which transformation algorithms to draw given its previous experience. Based on previous feedback, the system gradually learns which type of algorithms the user enjoys. It can then decide what type of line transformation is most appropriate in the current context.

In this reactive use case, the agent is working to make sense of the user's overall preferences through feedback, which form its experience. However, this process is reactive rather than interactive, meaning that the agent does not generate actions to test a dynamic model. Rather, it is finding what it believes is the optimal reaction to the current context. However, given this limitation, it is still important to note that novel and unexpected reactions do emerge. Additionally, the combinations of previous contributions with the current contribution can add more emergent visual elements.

These types of algorithms are not fully embodied because the agent is not perceiving the entire canvas and factoring in the current lines of the drawing when deciding what to draw. The decision is based on previous knowledge of the users preferences as well as the type of contribution the user made. If the agent incorporated the existing elements of the canvas and modified its contribution in the course of drawing based on evaluating how it affected surrounding regions, it would be more fully embodied. Using the reactive paradigm, the agent is learning from its experience interacting with the user, but the interaction history, i.e. the actual lines on the canvas, are not a part of its experience.

5.4.8 Discussion

Given its prominence in evaluating the agents behavior, spatial awareness seems to be a foundational skill for a co-creative drawing partner. One method for achieving spatial awareness is by constraining the agents learning mechanisms to particular regions of the artwork, i.e. executing certain types of drawing behaviors in certain regions based on user feedback. Since drawings develop over time, the perceptual logic in each region is also subject to change as regions grow, transform, and potentially connect with other regions. Given this dynamism, the systems learning algorithms should be temporally sensitive as well. For example, when users re-visit previously established regions, they may be trying to accomplish much different tasks, which could drastically change the perceptual logic that is appropriate. Considering both space and time in the learning algorithms should therefore improve the agents ability to coordinate with users and engage in participatory sense-making.

Our findings indicated that users did not fully understand how using voting and the creativity slider affected the behavior of the system. This could be mitigated by providing more explicit feedback about how their vote affected the agent's knowledge and drawing behaviors. In addition to describing how votes affect the creativity of the system, our findings indicated that users would benefit by reducing the ambiguity of binary feedback. Disambiguating user feedback could include providing a more continuous evaluation scale (versus the current binary like/dislike), as well as categories of feedback, such as providing independent feedback on the location, style, and content of the agents drawing contribution. In both conditions, participants used repetition to try to train and reinforce behaviors in the agent, but this type of implicit feedback was not registered by the co-creative agent. By analyzing the relationship of the user's input lines for a given turn, it might be possible to classify what type of contribution the user is making, i.e. defining a pattern through repetition, drawing a complete shape, or beginning a joint activity, such as coloring. All these modes of teaching emerged during the study, but the agent's machine learning architecture was not specifically designed to learn in that manner. Additionally, other implicit cues observed during interaction could serve as feedback for the system, such as when the user copied the agents contribution. Co-creative agents should take this type of mimicry as positive feedback indicating the early stages of a structural coupling that can facilitate participatory sense-making.

We also advocate a new wizard of oz user study design that includes a few new conditions: baseline, naturalistic wizard condition (i.e. collaboration decisions are unconstrained by how the agent works), as well as face-to-face collaboration. These new conditions would provide a spectrum of expressivity that would result in several different types of SM curves to compare these conditions in depth. Face-to-face collaboration allows the full range of human feedback (e.g. language, gesture, facial expressions, paper movement, etc.). Naturalistic collaboration with a wizard removes all human feedback but still allows the expert artist to collaborate in a flexible and naturalistic way without any channels of communication. The agent condition would remain the same. Introducing a baseline condition would help understand how the participant tends to engage in drawing activities without any intervention. With this experimental setup, there would be multiple levels of feedback fidelity that produce progressively less rich (i.e. varied) SM curves and we predict this experimental design would improve the analytic power of such a study in the future.

5.5 Conclusions

This chapter described several different user evaluations at different stages of development of the Drawing Apprentice prototype, which is a co-creative drawing partner. The system was designed to improvise and collaborate with users in real time on a shared artwork. The conceptual framework of participatory sense-making was adapted to identify important elements of the user experience and interaction design of co-creative systems. Since creativity and co-creative systems are notoriously difficult to evaluate, several types of studies and evaluations were employed, including formal user studies investigating the different between human and human-computer collaboration, user studies investigating how users interpret the interface and interaction design of the system, informal demonstration-based evaluations during art exhibits and public demonstrations, as well as expert evaluations of artworks produced with the system. We identified critical metrics users employ to evaluate whether the systems contributions made sense and describe how participants worked to train and provide feedback to the system to help coordinate their interactions. We leveraged these findings to propose design recommendations for co-creative agents. These findings help answer the second and third research question of this thesis that focus on quantifying interaction dynamics of collaboration and technical approaches to facilitate human-like collaboration.

CHAPTER VI

PARTICIPATORY SENSE MAKING IN CREATIVE IMPROVISATION

6.1 Summary

This chapter presents the results of an empirical study of 32 adult dyads (i.e. groups of two people) engaged in pretend play. Understanding the highly improvisational domain of pretend play can help identify key elements of coordination and collaboration relevant to co-creation in general. Our analysis indicates that participatory sense-making plays a key role in the success of pretend play sessions. We use the cognitive science theory of enaction as a theoretical lens to analyze the empirical data given its robust conceptual framework for describing participatory sense-making. We present here five enactive characteristics of pretend play that appear to be necessary and sufficient for the emergence and maintenance of successful pretend play mental preparation, meaning building, narrative enaction, narrative deepening, and flow maintenance. This enactive characterization is used to propose a computational model of pretend play that exands on the enactive model of creativity presented in the Chapter 3. Together, this work creates a theoretical framework to understand improvisational creativity as well as collaboration in open-ended creative domain that can be used to inform the design of an agent capable of playful co-creative collaboration with human users.

6.2 Introduction

Play is a fundamental aspect of human existence. Although play predates any concept of human culture or society [93, 132] animals engage in play as children and adults without any formal cultural context it is an important part of the human condition within familial and social groups. Play serves to strengthen social ties within groups, increase affect between individuals, and allow meaningful learning and practice at creative problem solving [14]. While play has been categorized by multiple efforts, it has yet to be formally understood in terms of the processes and actions participants execute to create a story world together, make stories, and establish shared meaning. Studying the fine grained behaviors of individuals engaged in pretend play can therefore inform us both about play at a deeper level as well as provide insight into how to formally represent such behaviors in computational systems. These formal representations can in turn help the design of various technologies to support, facilitate, and teach playful behavior.

This chapter describes our current efforts to characterize successful playful behavior between adult dyads (groups of two people) with an aim towards informing intelligent agents that are capable of playing with human collaborators for entertainment, learning, and play therapy. Our current specific focus is on studying the socio-cognitive capabilities involved in third person pretend play between adult dyads (i.e. play between two participants who physically control objects and characters) [147, 166, 188]. We present a theory of pretend play based on our empirical observations viewed through the lens of the enactive theory of cognition.

The enactive approach in cognitive science emphasizes the social and intersubjective nature of human understanding [135]. While our analysis may have employed other cognitive theories, such as embodiment, distributed cognition, situated action, social cognition, or information processing, enaction provides a framework that unifies elements of each of these approaches together, which helps provide a systemic perspective of pretend play. In particular, enaction emphasizes the role that emergent and dynamic social coordination plays in guiding and facilitating perception and action [64, 162]. We leverage the robust conceptual framework and vocabulary of enaction to formally represent participatory sense-making in the domain of pretend play.

Enactive cognition explains interaction dynamics, striving primarily to understand how perception and action are coordinated with the environment and other agents in that environment through emergent and continuous interaction known as structural coupling (or simply coupling). In this theory, stable relationships between perception and action characterize co-constructed meaning in the environment (i.e. the rules of the game that help guide behavior and frame expectations to facilitate successful interaction) [66].

In his work detailing the enaction paradigm, Vernon [174] describes sense-making as the process by which "emergent knowledge is generated by the system itself [as] it captures some regularity and lawfulness in the interactions of the system, i.e. its experience." Our empirical study of play, as described in this paper, suggests that the primary process or mechanism that drives dyadic pretend play can be described as participatory sense-making (multiple agents engaged in coordinated sense-making), per the enactive theory of cognition [66].

We contend that successful pretend play requires players that are willing to a) co-construct shared meaning, b) enact a narrative based on that shared meaning, and c) deepen the narrative in a coordinated manner to maintain the flow of the emergent play experience. There are many communication, interaction, and cognitive strategies and processes recruited in successful pretend play, but our primary contention is that participatory sense-making is the fundamental phenomenon that gives rise to successful dyadic adult pretend play.

6.3 Experimental Design

We conducted an observational experiment to investigate pretend play during which we recruited adult dyads (i.e. groups of two) to play together in different conditions. Overall, 32 dyads were recruited, with a total of 64 participants. Recruiting advertisements specified to bring a partner to the study (i.e. participants were not playing with strangers). Participants were recruited from the student population of the Georgia Institute of Technology (age range 18-24; n=33 male, n=31 female). Of the 32 pairs, 16 consisted of male/female pairings, and the other 16 were pairings of the same gender (male/male, female/female).

After warmup improvisational activities, participants completed two pretend play sessions lasting five minutes each, which were recorded, resulting in 64 play sessions. The play sessions took place on a large play-mat laid out over tables to allow players to stand while playing. Toys were kept in a box on the edge of the table containing primary-colored foam blocks and a varied selection of toys, such as those shown in Figure 26 and Figure 27.

Participants were randomly assigned one of four scenario prompts to guide their play: Drag Race, Car-Smash-A-Thon, Monsters Attack, and Zoo Visit. During the first play session both participants were given the same (randomly-selected) prompt to guide their play, while during the second session, their prompts differed (referred to as session A and B in data analysis, respectively). Half of the 32 dyads groups were asked not to talk (sound effects were permitted) during their sessions in order to investigate the effect of verbal communication on pretend play. In all conditions (talking and non-talking), participants were encouraged to play together and find a way to use both of their prompts in the same play story. After each session, we administered a retrospective protocol analysis during which participants were shown their filmed play session and asked to describe their motivation, intention, and general thoughts on the actions they took during the play session.

Figure 26: Experiment setup of toys and play mat with two participants from the adult dyad study.

Figure 27: Main character toys from the pretend play study.

6.4 Data Analysis Method

Since relatively little is formally known about the sociocognitive processes of pretend play, we designed our data analysis method as an exploratory investigation to characterize playful behavior. We utilized a Grounded Theory [74] approach to the data analysis that began by reviewing the video records from the pretend play studies and coding the data to identify prominent concepts and categories. Initially, we framed our analysis purely in terms of identifying all the observable behaviors involved in human dyadic pretend play to embrace the bottom-up, data-driven approach of grounded theory.

Through gradual iteration, we devised a categorization and coding scheme that described actions and related concepts. Early categories included: Player, Object Type, Object Role, Play Action, Communication, Narrative Development, and Milestones. Within each category, there were often many nuances and subcategories. As our analysis continued, it became clear that the dynamic and flowing nature of participant interactions could not be explained by any one action or combination of actions. The success of play appeared to be correlated to some emergent property of multiple factors. After comparing our empirical play data to the processes described in enactive literature on sense-making, we hypothesized that pretend play and participatory sense-making feature a similar process of social coordination utilizing the history of interactions, negotiated meaning, and feedback from verbal and non-verbal communication. With this observation and insight from the initial coding set, we iterated on our coding scheme once more by leveraging the concepts of participatory sense-making in enaction that help describe interaction dynamics.

We scoped our research question as a means of operationalizing our data-driven insights and reframed the investigation to ask: what are the minimal requirements to enable an agent to successfully play? To answer this question, we framed our analysis using concepts from the theory of enaction and focused primarily on a) continuously evolving interaction (rather than discrete actions and cognitive scripts) and b) different ways of coupling and coordinating interaction between agents to build meaning in a way that leads to successful play.

This type of analysis required an event level description of what types of perceptions and actions players used to make sense of the current interaction throughout the play session. This included a description of what actions the players performed and what analysts inferred they were trying to achieve with those actions given the current and historical context. To acquire this data, we performed an event level textual description of all the videos by carefully watching and transcribing an intentional description of what analysts inferred participants were trying to accomplish, a behavioral description of the how participants performed the actions to accomplish their intention, and an evaluation examining how this particular interaction related to the perceived success or failure of the play session.

6.5 Enactive Characterization of Pretend Play

Our data suggests that there are five critical ingredients required for two agents to successfully play: 1) Enter into a playful mindset, willing to engage in imagination; 2) Negotiate a set of rules and roles that constitute a nucleus activity; 3) Embody characters and interact through them in a shared narrative world; 4) Introduce creative actions and elements to make the narrative more interesting; 5) Ensure coordination by negotiating timely additions to the narrative. Each of these ingredients is described in detail below referring to empirical data from the play session, as it is helpful to describe the characteristic. The play sessions are numbered 1-32 and denoted with an A or B depending on whether they were the first or second play session of the experiment, respectively.

6.5.1 Prepare the Mind

Enter into a playful mindset to frame the interaction and set expectations. While pretend play typically comes easily to children, adults may feel self-conscious and perhaps even silly playing with toys and creating an imaginary story world. For play to be successful, participants should be open and willing to *suspend their disbelief* and work to fully immerse themselves in the narrative world. Preparation strategies observed in the data include taking on the persona of a character and beginning to interact with the environment through that character. Actions that signal a player is attempting to *embody* the persona of a character provide evidence of mental preparation. For example, participants often lowered their voice and moved more slowly when controlling large monster characters, such as Godzilla.

Players who failed to prepare themselves during the warm-up activities also tended to fail to immerse themselves in play, as was the case for Session 25. During Session 25, Player 1 appeared uninterested in playing, as evidenced by minimal participation in the warm-up games; that player attempted to gloss over each game by doing the bare minimum required to finish the game or let the timer run out. Based on our observation, this player was not open and willing to become immersed and play in an imaginary world. The data indicates that the more immersed players become, the easier it is to generate actions to perform, which can lead to more successful play (as shown in examples in the next subsection).

6.5.2 Build Meaning

Negotiate a set of rules and roles that constitute a nucleus activity and shape interaction. Players co-construct a new reality, a shared narrative world, by physically and conceptually structuring the environment in meaningful ways, taking control of characters, and providing details and specifications of characters that help enact the narrative [135]. Without a basic foundation of shared meaning, the participants do not know the 'rules of the game,' so to speak, and therefore cannot enact a narrative and successfully pretend play. We define this minimal seed of shared meaning as a nucleus activity, which is the most clearly defined and agreed upon elements of a story world and their most prototypical associations (i.e. prototype theory of categorization [140]). By definition, nucleus activities contain at least one role for each player and one rule to guide and shape interaction in some manner (see Figure 28).

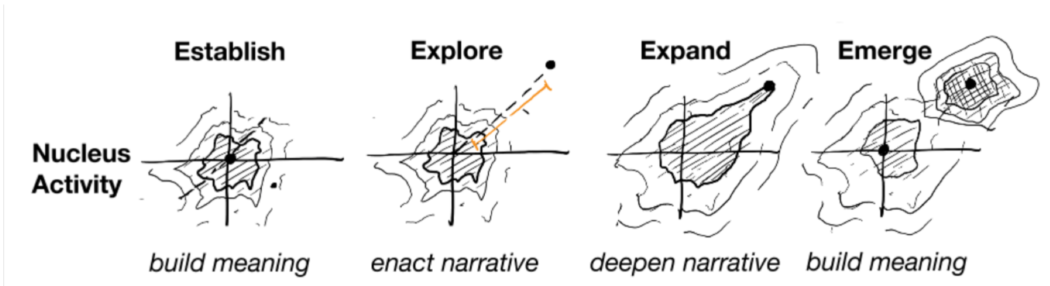


Figure 28: Depiction of nucleus activity growing through time.

The nucleus activity consists of a solidly negotiated core with peripheral concepts that are tangentially related for either participant. The strategies participants use to build meaning and co-construct nucleus activities vary drastically. The number of elements used to add onto the nucleus activity and build the narrative world, for example, does not seem to necessarily correlate with the success of the play session. Rather, the quality or depth of meaning attached to each of those elements influences success. Assigning more details to further specify the role of their characters facilitates a deeper character embodiment. As players become more deeply immersed in the narrative world, they subsequently interact more naturally through that character.

This finding demonstrates that the quality of meaning that is co-constructed and applied to elements in the play space is more influential that the number or type of elements used in the play session. Individuals that were obviously not immersed in the narrative tended to have less qualitatively meaningful elements in the narrative world, which suggests that preparing the mind and being consciously open and willing to immerse oneself in an imaginary world is correlated with the depth and complexity of meaning co-constructed in the narrative world.

6.5.3 Enact the Narrative

Embody characters and interact through them in a shared narrative world. Once a nucleus activity is well established, players perceive the real objects of the environment (i.e. blocks and toys) through a *perceptual logic* [31, 36] that filters perception with respect to the co-created meaning structures of their nucleus activity. Examples of perceptual logic that could account for interaction patterns in pretend play include character motivations, character play affordances, narrative trajectory, environmental constraints (e.g. the setting), and feedback (e.g. other players).

Actions are not generated solely from a narrative or cognitive script. Rather, actions emerge through embodying and taking on the persona of a character and performing actions that make sense for that particular character in that particular narrative world (which may happen to draw upon previously learned cognitive scripts). Character definitions, motivations, and tendencies are adjusted based on feedback from their play partner. Narrative is an emergent quality of pretend play that arises as players work together to make sense of their respective actions (both retroactively and proactively) in the context of meaning structures established thus far in the play session. We propose that this social coordination through interaction is a form of participatory sense-making and a key component of describing pretend play.

Participants generally agree upon the basic rules and roles of a nucleus activity, such as 'monsters fighting, but explore the search space of the nucleus activity and push its boundaries through the process of enacting the narrative. When participants disagreed, it was because there was a further specification that was assumed by one player given the agreed upon nucleus activity, but that assumption was not shared by the other player.

6.5.4 Deepen the Narrative

Introduce creative actions and elements to make the narrative more interesting. Purely enacting a basic narrative is engaging for a short period of time. To maintain creative engagement for an extended period of time, it seems necessary for players to add additional details and elements to the story world. This aspect of participatory sense-making no doubt has different strategies. We observed one strategy in particular that appears to be a recipe for success.

First, a nucleus activity is negotiated during initial setup. That nucleus activity can contain different amounts of complexity and detail. It can be negotiated using a variety of methods, but it minimally involves a definition of rules and roles. Those rules and roles have relevant knowledge associated with them, which should be considered as being included in a *shared conceptual search space* of the co-constructed nucleus activity. Each action players perform has a certain semantic distance (degree of relatedness between concepts) from the core of this nucleus activity. Actions that are further away from the core are defined as more creative.

Creative actions require more explicit forms of negotiation because they might fundamentally change the nucleus activity and narrative world based upon it. When distant creative actions are not successfully negotiated, *siloed play* may occur as each players mental model of the narrative world diverges. Successfully negotiating creative actions expands the core of the nucleus activity, as shown in nucleus activity expansion phase of Figure 4. Since the conceptual space of the nucleus activity is by definition a shared search space, its expansion increases the possibilities for relevant interactions, which tends to make it easier for individuals to play successfully.

Questions and actions that help clarify and add specificity to elements of the nucleus activity help to enact a narrative. For example, as players in 33A walked their character around the zoo, they questioned how the animals were caught, which created an opportunity to provide an interesting back-story. Player 1, as his Godzilla character asked, "How did you manage to catch this giant tiger?" Next, Player 2 responded with a witty retort, "With a lot of cat nip" When players rationalize their selections with respect to the nucleus activity, they tend to help make the narrative world more robust, interesting, and creative.

6.5.5 Maintain the Flow

Ensure coordination by negotiating timely additions to narrative. The creativity of participants and the actions they perform must be paired with the ability to maintain the flow of the play session through time. Successful sessions typically featured players that were attentive to their partner and strived to include them in a meaningful way. Depending on the demands of the situation, this can include subtle gestures, such as seeking feedback using eye contact. More active maintenance activities involve explicitly engaging their partner, such as directing actions and dialogue toward them, or asking their partner questions to prompt elaboration. Social skills, such as empathy are important here.

Good players maintain a healthy respect for the rules of the nucleus activity, and will defend actions that violate those rules in some way (while still remaining open to negotiation). When players take creative actions that could be classified in the distant periphery of the nucleus activity, sometimes negotiation is required to ensure the nucleus activity expands properly. Similar to how a good conversationalist knows when a topic is becoming stale, good players consciously maintain the flow of the play experience. Players engage in a coordinated dance of building on and subverting their partners intentions in the shared narrative world by modulating between enacting and deepening the narrative. This skill involves knowing when to add depth to the narrative world and how to include your partner in that process. Through time, creative activities expand nucleus activities into new domains that might require slightly more rule definition and specification, eventually forming an independent nucleus activity, as shown in emergent nucleus activity phase in Figure 29.

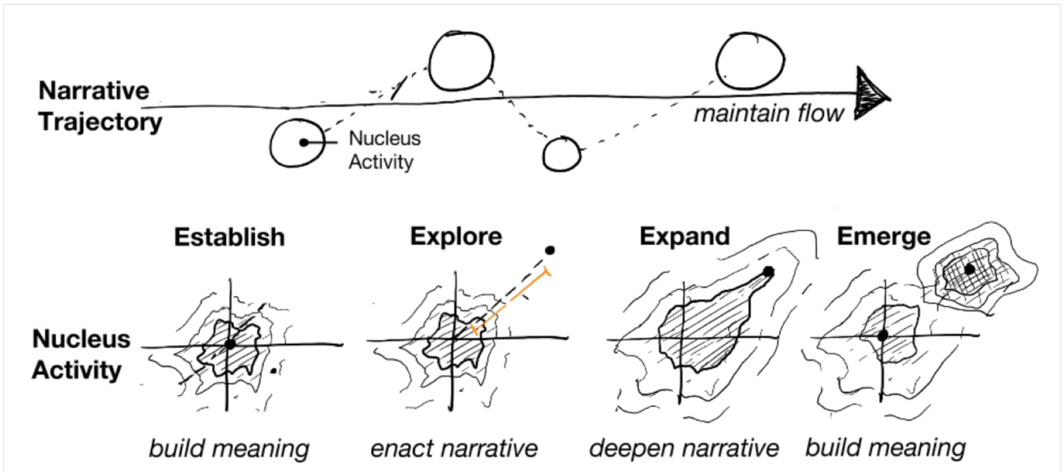


Figure 29: Depiction of narrative emerging from collection of nucleus activities through time.

Successful play sessions tend to have relatively well defined leader/follower roles that naturally switch over time as players come up with new ideas and strive to implement them in the story world. Oftentimes, the most successful play sessions involved players who handed off leadership to each other as their narratives progressed. Players that exhibit leadership in play tend to work to make sense of both their and their partner's play actions by developing a common thread tying together the various nucleus activities constructed throughout the play session, termed the *narrative trajectory* and shown in Figure 29.

6.6 Conclusions

This chapter reported on an empirical investigation into pretend play between adult dyads. We used the cognitive science theory of enaction as a lens to analyze our empirical data and developed an enactive characterization of pretend play. In particular, we propose five characteristics of play that all rely on participatory sense-making: preparing the mind, building meaning, enacting the narrative, deepening the narrative, and maintaining the flow of the play session. The enactive concept of participatory sense-making was proposed as the key mechanism of pretend play. We developed a novel graphical convention called sense-making curves to model and represent interaction dynamics over time. Our future work includes conducting another round of data analysis to plot sense-making curves for all the pretend play sessions. This data will help evaluate the predictions and hypotheses generated by our enactive characterization of pretend play.

CHAPTER VII

QUANTIFYING INTERACTION DYNAMICS

7.1 Summary

This chapter provides a detailed description of the proposed technique for quantifying interaction dynamics in open-ended creative improvisational collaboration. It begins by introducing the problem and describing the current techniques used in cognitive science to analyze interaction dynamics. Then, it proposes the existing free-energy theory of the brain as a universal means of quantifying cognitive states, i.e. whether free-energy is minimized or not. Sense-making is described in terms of free energy as the process by which free energy is gradually reduced in the brain by experimentally interacting with the environment to gradually increase the accuracy of the agents internal predictive model. The sense-making curve is described as a way to continuously record an agent's free energy and subsequently quantify their sense-making activities.

7.2 Introduction

Several metrics have been developed relating to creativity and technology. The creativity support index (CSI) is a psychometric survey instrument that measures the effectiveness of a creativity support tool for assisting users engaged in creative work [5]. Protocol analysis has been used as a method for evaluating and comparing how different user interfaces and input methodologies affect creative cognition [23]. Some aspects of the users creative process can be quantified by logging user data from creativity support tools, such as measuring the emergence of new ideas and ideation strategies [22], the novelty and surprise of designs [27], and the effectiveness of a creative collaboration [8]. Biometric sensors, such as EEGs, have also been employed to quantify in the moment creativity (ITMC) by classifying periods of heightened creativity based on physiological markers [4].

Due to the complexity and open-ended nature of creative activities, researchers generally employ a mixed-method approach of data triangulation that draws on multiple sources of data to analyze and evaluate how a technological intervention affects the creative process [4]. While these creativity research methods provide insight into an individuals creative process and tools utilized during creativity, evaluating creative collaboration presents unique challenges around understanding how collaborators coordinate in the moment to co-construct shared meaning throughout a creative collaboration.

The enactivist paradigm in cognitive science has made significant advancements in terms of understanding how meaning emerges through interaction, both by an individual agent and through social coordination [3][10]. These researchers propose a novel theoretical framework focused on the idea of sense-making, whereby an agent gradually casts a web of significance and meaning onto the world through interacting with the environment (and other agents within it) to determine meaningful regularities [32][34][33]. The theoretical framework of sense-making is conceptually robust, replete with a vocabulary and paradigmatic viewpoint for understanding cognition and interaction through the lens of socially emergent and dynamic meaning constructs [13]. However, there is a significant gap in the field regarding quantifying the interaction dynamics of sense-making during complex and open-ended activities, such as improvisational creative collaboration. In the enactivist literature, there are generally two approaches to quantifying interaction dynamics and sense-making: perceptual crossing and traditional qualitative analysis.

7.2.1 Perceptual Crossing Methodology

Perceptual crossing is a type of participatory sense-making and experimental apparatus recently introduced into the cognitive science literature to quantify interaction patterns directly within an artificially constrained virtual environment [2, 41, 42, 62]. Researchers studying perceptual crossing utilize a minimalist virtual environment in which two agents (either both human, both artificial agents, or a combination thereof) perform a spatial participatory sense-making task whereby they try to differentiate the motion of their partner from static objects using only tactile feedback. Each player moves their avatar across their respective screens and the participants receive a vibration if the avatar has crossed paths with something in the environment (whether it is object or human). In this virtual environment, all actions are restricted to spatial keyboard inputs and are thus easily observable and quantifiable. This approach reveals some of the underlying processes of participatory sense-making, but it is not applicable for understanding sense-making in more complex domains, such as open-ended and creative interactions.

7.2.2 Traditional Qualitative Analysis

The second approach for understanding interaction dynamics is traditional qualitative analysis, i.e. qualitatively coding observational video data of complex social activities to interpret the types of actions and strategies recruited during participatory sense-making [97, 159]. This approach is widely used in dialog and conversation analysis [160] and analyzing turn taking dynamics, such as leader/follower strategies and topic shifts throughout the interaction [7]. These investigations typically yield a thick cognitive ethnographic description of the factors that influenced participatory sense-making, but they are difficult to quantify and analyze statistically.

Qualitative analysis employing event-based coding practices are also common in analyzing and evaluating open-ended creatiity, such as artistic and design creativity. For example, Maher (2006) adopted this approach when evaluating collaboration practices of designers in a virtual environment, segmenting data into discrete events based on behavioral markers [108] following a similar approach established earlier by [49]. Yokokochi and Okada [185] and Mace [104] also employ event-based qualitative analysis to understand the open-ended artistic creative process. Interestingly, Yokokochi and Okada's analysis attends to the behavioral markers of pausing and body-repositioning (i.e. stepping back from the artwork) as important events in the coding scheme similar to the current approach. These approaches yield powerful and imformative descriptions about the number of times events occured and even the order in which these events occur, but they still do not provide continuous data that can be analyzed using continuous mathematical functions to quantify the fine-grained temporal dynamics of interaction.

7.2.3 Creative Sense-Making Analysis

Creative sense-making bridges these two general methodological approaches by creating a simplified qualitative coding scheme focused on sense-making that lends itself to quantification and computational analysis. In this approach, each participant's action is categorized according to its functional role in sense-making. We employ the free-energy principle of the brain [59–61] to develop a quantitative descriptor for these different states. The free-energy principle claims that cognition continually strives to reduce surprise, i.e. cognition strives to create a dynamic and generative mental model that makes the environment more predictable [60, 150]. A few critical definitions will help elucidate the theory of free energy.

- **Generative model:** "or forward model is a probabilistic mapping from causes to observed consequences (data). It is usually specified in terms of the likelihood of getting some data given their causes (parameters of a model) and priors on the parameters" [59].

item textbfSurprise to agent: this occurs when a cognitive agent has developed a generative model of interaction and anticipates certain sequences of data using that model, but the data from the environment violates the expectation of the agent.

- **Free energy:** "Free energy [is] an information theory measure that bounds or limits (by being greater than) the surprise on sampling some data, given a generative model" [60].

- **Free-energy principle:** "The free-energy principle says that any self-organizing system that is at equilibrium with its environment must minimize its free energy." [60].

Similar to sense-making, the free-energy principle claims that humans interact with the environment, through both active perception and action, to improve the accuracy of their generative model of the environment, thereby reducing *free-energy* [60]. When free energy is minimized, actions are generated fluidly and with ease, allowing agents to directly perceive affordances and meaning constructs in the environment that increase the order and predictability of interactions. When free energy increases, i.e. the cognitive agent becomes surprised, perception and action are utilized to help increase the accuracy, or recognition density, of the generative model of the environment, i.e. making the predictions of the generative model more closely match the true conditions of the situation.

Combining the free-energy principle with the conceptual framework of sense-making enables a new method for quantifying interaction dynamics based on the relative free energy of a cognitive agent through time, as determined through behavioral markers. In the proposed approach, when agents do not have a robust generative model of the environment, they have to engage in a process of sense-making, which costs physical and mental energy. This type of sense-making can be viewed as an investment of physical and mental energy that has the potential to reduce free energy, in the long run, i.e. improving the predictive power of the cognitive agents generative mental model. Thus, sense-making is formally defined here as the process whereby a cognitive system gradually minimizes free-energy by experimentally interacting with the environment to build and refine a more optimal generative mental model of that environment.

Within the context of this framework, we propose two basic states of cognition that have clearly delineated yet interrelated functional roles. Borrowing terminology from Glenberg [19], these cognitive modes are referred to as: clamped and unclamped cognition. The concept of clamping represents a cognitive mechanism that helps the agent balance exploration versus exploitation in order to learn a better, more accurate, generative model (i.e. make sense of the environment).

- **Clamped Cognition:** The process of maintaining or slightly refining the selected generative model assuming that it is the most accurate representation of the environment. It generally occurs after making sense of a task or activity. Behavioral markers include fluid interactions with minimal hesitation (e.g. embodied play actions, fluid drawing actions).

- **Unclamped Cognition:** The process of changing or replacing the generative model by exploring the environment from different perspectives. It generally occurs during task onset and after surprises during the task. Behavioral markers include hesitation (e.g. eyes closed, confused look) and physically experimenting with the environment and viewpoint of the environment (e.g. futzing, inspecting).

We devised a set of behavioral markers to detect clamped and unclamped cognition through a qualitative coding procedure of video data. For example, in the context of pretend play, a participant that is embodying a play character and performing fluid play actions inside the narrative world (i.e. diegetic actions) is considered clamped. Conversely, if a player is hesitating, unsure, pausing, or otherwise disengaged, this is a sign they may be actively processing and working to make sense of the situation to determine an effective strategy for moving forward, which signifies an unclamped state. Pauses and hesitations are also used in conversation analysis and analyzing interaction dynamics [2]. When players restructure or build additional meaning into the environment explicitly (i.e. extra-diegetic actions), they are actively making sense of the situation in an unclamped manner.

In creative interaction, meaning is continually shifting and evolving due to the participatory nature of improvisational collaboration [28]. However, empirical evidence [7] suggests that semi-stable (yet dynamic) meaning structures emerge to produce a steady state whereby both participants have a relatively robust predictive model that enables them to :interact fluidly in a situated manner, without much explicit sense-making outside of fine-grained coordination. In our work, these stable units of meaning are referred to as nucleus activities that can grow through time as additional layers of meaning are added [7]. Nucleus activities arise when both participants are clamped on a relatively stable meaning structure that can guide their interaction through time and dynamically grow (i.e. adding new layers of meaning that slightly change the 'rules of the game'). Conversely, participants can be confused or uncertain about what type of action to take or how to interpret their collaborators actions due to a sparse predictive model, which would be characterized as unclamped. This type of clamping framework extends the idea behind shared meaning established in the literature about shared mental models [20][6] by taking into account the dynamic and evolutionary nature of meaning construction in open-ended creative interactions.

Within the unclamped category, there are two further distinctions that can be made to increase the granularity and explanatory power of our proposed coding technique. Unclamped actions that are meant to reduce free energy through sense-making can be either perceptually-based or physical-based. Perceptual-based sense-making relates to refining the brain's predictive model, which subsequently changes how features in the environment are perceived and interpreted. Since these processes are happening internally, they cannot be directly observed, but individuals experiencing this cognitive state display indirect behavioral markers, such as pausing, hesitating, contemplating, and looking confused, i.e. thinking. Physical sense-making, on the other hand, changes the structure of the environment by manipulating and modifying the environment, or by moving the body and altering what information is available to the senses, i.e. thinking by doing.

These two different paths of reducing free energy represent opposites that mirror the afferent and efferent flows of sensory information in a cognitive system. Perceptual-based sense-making processes change the afferent flow of predictions being generated by the brain and projected to the body (via the nervous system) by thinking and re-evaluating the situation internally. Physical sense-making processes change the efferent flow of incoming sensory information available to the agent (to generate predictions about) by taking actions that either change the environment (through manipulation) or one's vantage point of it. Thus, our proposed approach makes two chief categorical distinctions: clamped and unclamped cognition, and within the unclamped category, there is a further distinction between perceptual sense-making and physical sense-making.

- **Perceptual sense-making:** the cognitive agent is working to internally improve recognition density of its generative mental model, i.e. thinking. Behavioral markers include: hesitation, eyes closed, confusion, and task disengagement in general.

- **Physical sense-making:** the cognitive agent is exploring the environment through interaction to decrease disorder in the environment and increase the recognition density of its generative mental model, i.e. thinking by doing. Behavioral markers include experimentally manipulating resources in the environment and re-positioning the body to change available sensory data.

There are also degrees of each unclamped category. For example, agents can pause their actions to wait for their partner to take their turn (partial perceptual unclamp), or they can be completely confused and disengaged from the play session (full perceptual unclamp). Complete confusion and disengagement should be considered more unclamped than attentively waiting for a partner to act. Conversely, in terms of physical sense-making, searching for new resources to introduce into the play session would be more unclamped than slightly rearranging or restructuring elements already in the play space. Physical sense-making and perceptual sense-making represent two different ends of a spectrum of sense-making strategies. As a result, this categorical distinction is reflected by assigning one of these types of sense-making processes as a positive value and the other a negative value. The decision of which category receives a positive vs. negative sign is arbitrary. The important delineation here is providing more granular data about the type of unclamp in which the participant is engaged. The magnitude of the numerical value assigned to the unclamped action is determined by assessing the degree of unclamp based on behavioral markers. To help systematize the coding process for unclamped actions, we use only two levels to describe unclamped actions with a value of .5 or 1 (in both the positive and negative direction).

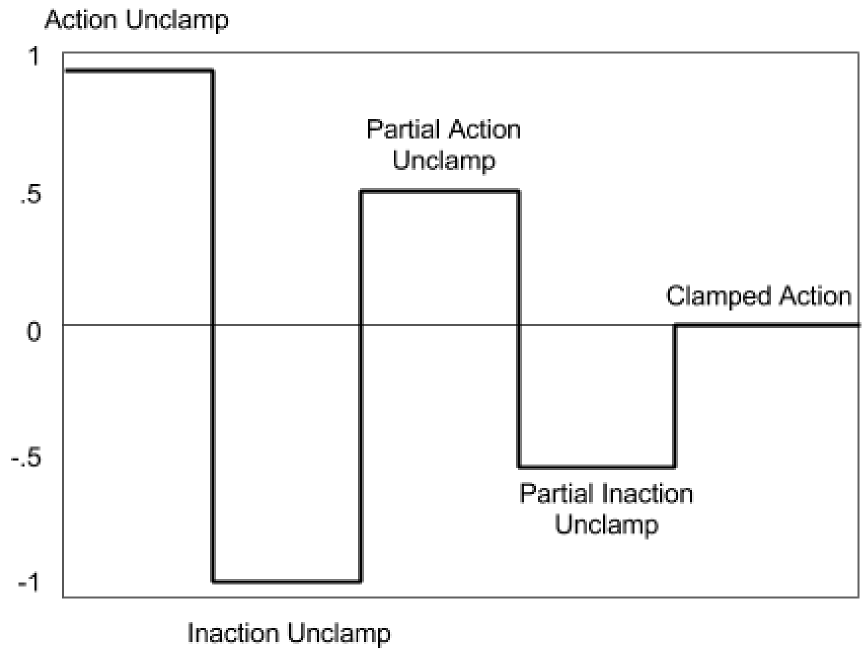


Figure 30: Generic sense-making curve demonstrating clamped/unclamped states

When plotted along an x-axis representing time, these numerical values create what we refer to as the sense-making curve (as shown in Figure 30).This curve quantifies what types of actions each individual participant was engaged in throughout the course of the experiment. These curves can be mathematically analyzed and combined to quantify the interaction dynamics between the players.

To summarize, we propose a new approach to formally model the interaction dynamics of participatory sense-making. We use the qualitative video coding conventions of the sense-making curve to translate human interaction dynamics and patterns of interaction into a machine readable format. In the next sections, we describe the web-based tool that was developed to perform this coding procedure and the analysis technique for combining sense-making curves and classifying different types participatory sense-making and styles of collaboration from sense-making curves.

7.3 Method

The prototype sense-making curve tool (shown in Figure 31) is a web-based qualitative video coding environment that utilizes videos uploaded to YouTube by the researcher. These videos can be set to private or unlisted on the YouTube platform to prevent public consumption of research data (the private setting is strongly advised to keep data password encrypted). YouTube provides a free online resource to store videos. It essentially functions as a storage platform for our web-based coding tool.

Figure 31: Web-Based Sense-Making Curve Tool

To code a video, the user pastes the URL address from the YouTube video into the setup page of the system (not pictured). Then, the system finds the video on YouTube and embeds that video into the coding environment shown in Figure 31. To apply the qualitative codes to the video, the researcher uses the up and down arrow keys to increase or decrease the current coding value, which is visualized by a draggable code selector slider on the right side of the screen. The system samples the value of the slider every 250ms and records that data point in an array list. Therefore, the value the code selector is currently resting at is the value that will be assigned to the time period that is currently being viewed in the embedded video window. The code the analyst selects represents the cognitive state of the individual (i.e. where they fall on the clamped/unclamped spectrum) during that time period.

The most recent values from the array list (relative to the current point in the video that is being played) are visualized in the recent data display to the right of the embedded video. When the user scrubs the video to different timepoints, the data display and coding value also move to that point in the array, allowing researchers to revisit different parts of the video to re-apply codes.

The embedded YouTube video has some of the original functionality from the YouTube platform, such as slowing down and speeding up the video playback speed. Researchers using this tool typically slow the videos down to half speed and code one participant at a time. The analyst re-watches the video to code each participant. Once an analyst was oriented to this approach, the overall code application time per video (including both participants) was approximately 4 to 1, meaning it took about 4 minutes to code every 1 minute of video data. A 4 to 1 time-efficiency ratio for qualitative video coding, and a 2 to 1 ratio for coding individual participants within an interaction makes coding large quantities of interaction videos more feasible.

It is important to point out that while the sense-making curve might facilitate rapid and reliable qualitative coding, the granularity of analysis between the sense-making curve and typical event-based coding techniques is much different. The sense-making curve provides continuous and targeted data about interaction dynamics, but it does not identify individual actions, i.e. the content of the actions. In future versions of this tool, we plan to allow tagging the curve with different events to combine the power of event-based coding with our proposed continuous coding technique.

The numbers representing each code on the 1 to -1 scale (according to the clamped/ unclamped convention) were each mapped to 0-4 values at the request of coders (to reduce time switching between common states in a given domain). As long as each

Table 5: Table showing code mappings between what codes analysts apply and how they map to the creatie sense-making theory

Clamp/Unclamp Value	Classification and Description	Code Application
1	**Full unclamped - action:** gathering resources	1
.5	**Partial unclamp - action:** building and changing environment	2
0	**Clamped - action:** fluid and embodied play actions	4
-.5	**Partial unclamp- inaction:** hesitation, attentively waiting	3
-1	**Full unclamp- inaction:** disengaged, distracted	0

state is consistently coded through time and the mappings to their placement on the 1 to -1 scale are stored, the numbers used in the user interface of the tool can change. This change in scale creates an interesting level of obfuscation between the analyst and the predictions of the theory and cognitive framework since the curve they are currently seeing visualized in the tool is not necessarily the final form of the curve used in the analysis (after the value transductions). It remains to be seen whether the tool is most reliable when used with the 1 to -1 scale or a 0 - 4 scale in the interface.

The data analyst's role in this context is to decide which of the five states participants are in based on behavioral markers. The design of the behavioral markers is determined by understanding what types of actions in a particular domain should be classified as clamped/unclamped and the degree of this association. See Appendix A for a breakdown of the coding scheme used for the pretend play video data analysis and Appendix B for its extension to the Drawing Apprentice user study analysis. A rough approximation is described in Table 5.

The coding tool and technique was used on both the pretend play dataset and the Drawing Apprentice user study data. For pretend play, three coders analyzed 64 videos, resulting in a total of 128 sense-making curves produced. These three

Table 6: Interpration of Fleiss' Kappa used for establishing inter-rater reliability.

Value	**Interpretation**
<0	Poor Agreement
0.01 - .020	Slight Agreement
.021 - 0.40	Fair Agreement
0.41 - 0.60	Moderate Agreement
0.61 - 0.80	Sustantial Agreement
0.81 - 1.00	Almost Perfect Agreement

Table 7: Table showing inter-rater reliability (IRR) for the sense-making curve coding technique applied to pretend play data and user study data.

Data Set	**Analysts**	**Codes Compared**	**Technique**	**IRR Score**	**Interpretation** [57]
Pretend Play	3	19,990	Fleiss' Kappa	.71	Substantial Agreement
User Study	2	362,560	Kripendorff's Alpha	.76	Substantial Agreement

coders were able to achieve an inter-rater reliability (IRR) score of .71 (substantial agreement according to [57]) using Fleis' Kappa (shown in Table 7). To calculate this score, an individual play session was analyzed by all the coders and the Fleiss' Kappa for the coded values of the left and right players in the session was averaged. This Fleiss' Kappa value compared the reliability of 19,990 total code applications between the analysts on the pretend play dataset. For the Drawing Apprentice study, two coders analyzed 12 longer user study videos resulting in 48 total sense-making curves between the two coders with an overall inter-rater reliability of .76 (substantial agreement according to [57]) using Kripendorff's alpha (for two coders). For this dataset, the IRR value for coding the participant (.64) and co-creative agent (.86) were averaged to yield the .76 value for Kripendorff's alpha over the entire dataset, which consisted of 362,560 code applications (shown in Table 7).

To achieve these reliability measures, two different teams of analysts coded the sample videos from each domain, tested their reliability, and looked for points where their coding diverged. Then, they discussed possible reasons for the divergence and added more detailed behavioral markers to the coding scheme to help ensure future reliability. After several rounds of refining the behavioral markers of the coding scheme, analysts were able to achieve substantial agreement over thousands of code applications. In most instances in the literature, these inter-rater reliability measures are not being used on datasets with sizes this large, or with continuous temporal data. For that reason, it may be beneficial to explore additional inter-rater reliability measures to pair with the Kripendorff's alpha and Fleiss' Kappa to help evaluate inter-rater agreement through time with increased precision.

Performing the numerical transduction on the data (from the 0-4 scale to 1 to -1 scale shown in Table 5) before performing these reliability measures may also yield insights into the productivity of this method. For example, once the data is represented in the 1 to -1 convention, it can be bucketed into three categories of positive, neutral, or negative. While this reduces the overall granularity with which the data is being compared, it provides a better indicator as to whether two coders were merely off in measuring the magnitude of the unclamp versus a completely different direction.

7.4 Limitations of Sense-Making Curve Coding Technique

The creative sense-making analysis technique presented in this chapter is a new data collection and analysis method meant to provide more quantifiable and fine-grained data to analyze co-creation and evaluate co-creative systems. The initial analysis demonstrates the types of questions this new framework enables us to ask. It also demonstrates some of the ways in which different types of collaboration strategies and styles can be quantitatively classified. While the results presented here provide an initial indication as to the utility the proposed method offers, this thesis was not focused primarily on formally evaluating the creative sense-making methodology. Instead, CSM was utilized as one method to help evaluate the Drawing Apprentice co-creative drawing partner. In that role, the CSM analysis helped provide insight about the behaviors of users throughout the study as well as how future studies might be designed to take full advantage of this new method.

Without full validation, the findings from the CSM curve analysis are still preliminary. There are a few concrete activities that will help validate this technique in the future. The primary concern with evaluating the CSM methodology is determining whether the classifications the technique generates (i.e. clamped/unclamped, mutual clamp, dominant clamp, etc.) actually correspond to those events in reality, i.e. does the theory correspond to the ground truth. In order to fully evaluate the ground truth of the CSM classifications, it is necessary to refer to the retrospective protocol analysis to determine whether the remarks the participant made during a given time period correspond to the predicted cognitive state predicted by the theory. However, since the CSM analysis method was developed after the Drawing Apprentice study was designed and executed, the retrospective protocol did not focus on soliciting information explicitly related to the cognitive states later identified by the CSM framework.

Another technique to establish ground truth using the CSM approach is to have researchers help participants produce their own sense-making curve describing their creative process after the study task is completed. Instead of doing a traditional retrospective protocol, the researcher can use the sense-making curve as a boundary object [xxx] to help facilitate precise descriptions of the participants cognitive state. The sense-making curve produced by the participant is valuable, but their targeted commentary about the curve is the data that will be extremely helpful when evaluating the efficacy of the CSM framework.

In addition to this limitation, there are procedural limitations arising from the process by which analysts apply codes in the CSM approach. In this approach, there is an inherent limitation in terms of the delay between an action onset and when a coder determines to move the coding slider to that corresponding value (since this is based on the analysts reaction time and subjective judgement). Depending on each coder's style, this delay may time-shift or change the granularity of certain events in the data, while still registering the overall deviation that is important later in the analysis. This limitation could drastically reduce inter-rater reliability while having an insignificant impact on the comparison of the actual sense-making curves and creative trajectories (as described below).

When users switch between codes using this tool, it is possible (and likely) that the system will sample data points during the process of moving from one code to another. This results in the system recording values between the current code and the target code. The number of values recorded between code switches increases depending on how far away the target code is from the current code since the user employs the up/down arrow keys to move between codes. This limitation of the tool creates some potential noise in the data as values are being recorded that were not intentionally coded by the analysts. However, this source of noise does not appear to significantly affect the overall reliability of the technique. That being said, one method to avoid this limitation is to use the number keys to control mode changes rather than the arrow keys.

7.5 Classifying Sense-Making Categories and Trends

The discussion below is based on analyzing sense making curves coded from raw video footage of pretend play sessions (and other similarly open-ended collaborative domains). From this point, we assume we have valid sense-making curve data and perform calculations upon that dataset. However, to begin the discussion we use a simplified sense-making curve for represenational purposes to help establish the sequence of mathematical operations visually. In the next section, we apply these techniques to an actual data set to further explicate the technique. The mathematical operations used in the analysis of sense-making curves are as follows.

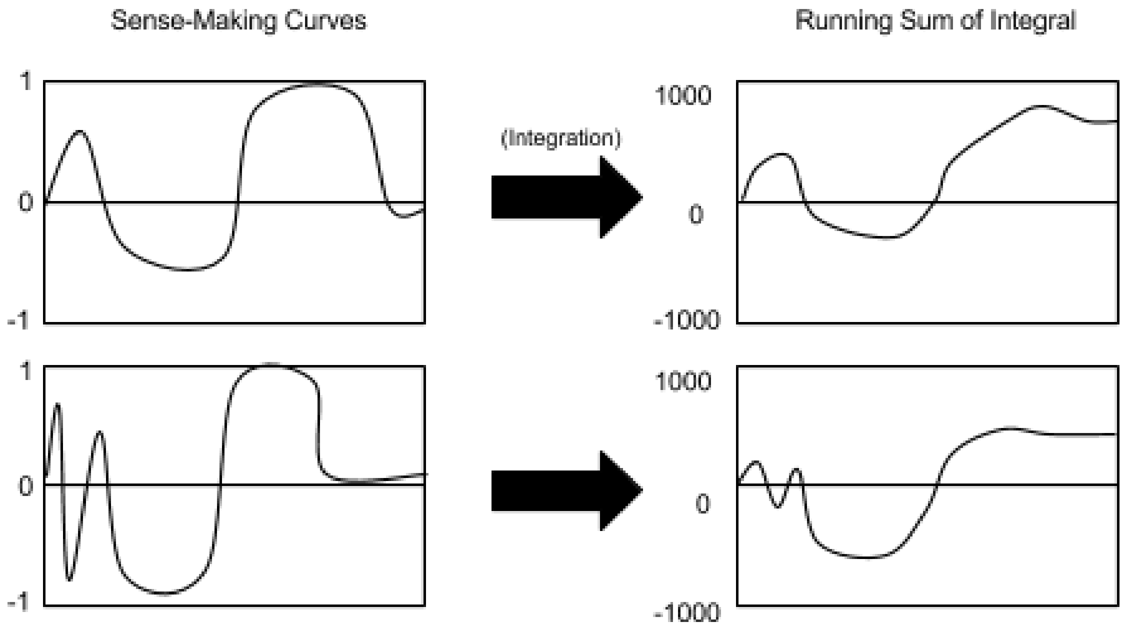


Figure 32: Transforming the sense-making curve into a running sum of the integral.

The starting point for the analysis is two sense-making curves collected from coding video data of a play session or creative collaboration (one curve for each player in the interaction). Both curves are integrated to create two new datasets with the running total of the integration for each curve. Finding the area under the curve represents how much unclamped action, thinking, and clamped play a player has engaged in up to that point in the play session (as shown in Figure 32).

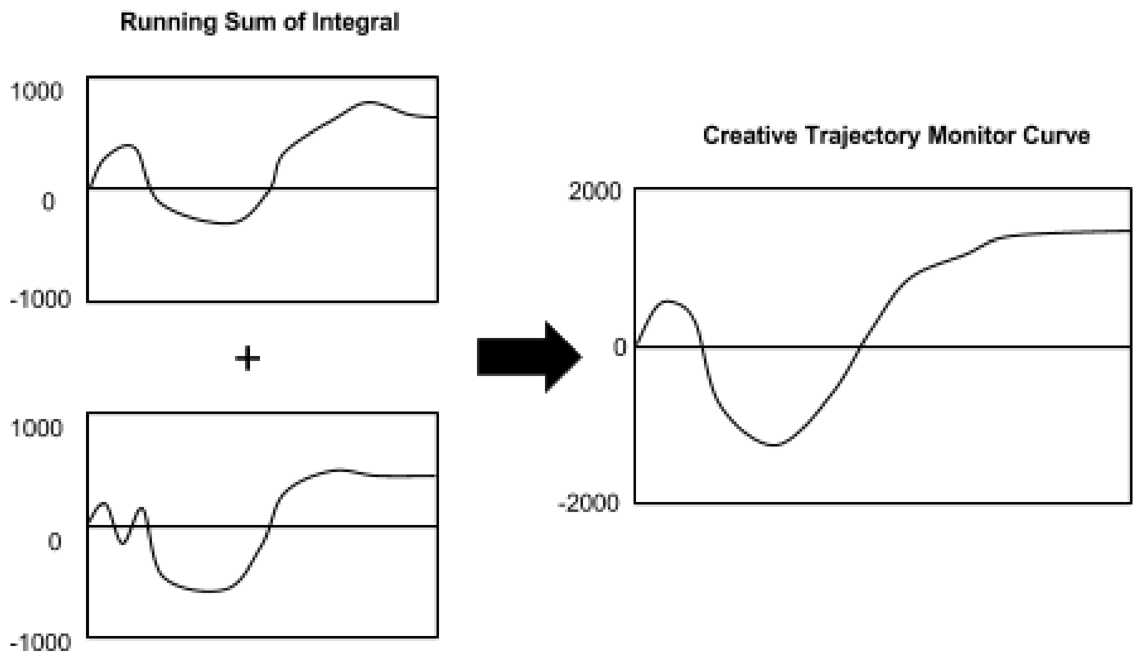


Figure 33: Summed total of the running sum integral yields a combined creative trajectory

Next, the integrals are summed to yield a cumulative value describing how the area under each player's curve related to each other through time called the creative trajectory (shown in Figure 33).

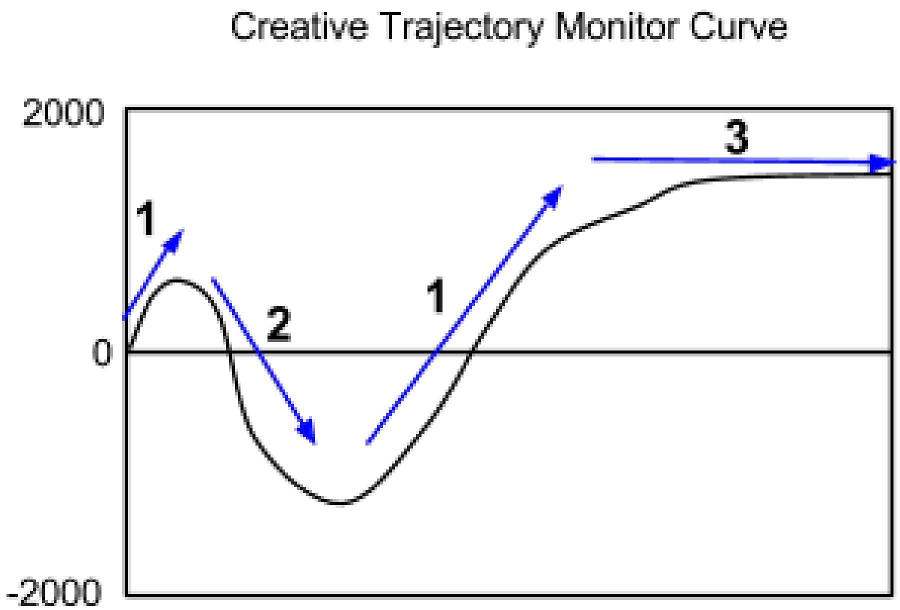


Figure 34: Defining the collaborative momentum with the creative trajectory curve

Finally, we measure the rate of change of the creative trajectory, i.e. its derivative, to detect a value that can be termed the *collaborative momentum*, or the general and combined direction that the collaboration is headed in at a given point in time, e.g. more creative possibilities, less, or holding steady. The collaborative momentum is the average rate of change of the collective running sum of two players' sense-making curves (shown in Figure 34).

7.6 Interaction Dynamic Framework

This section begins to sketch a preliminary framework for quantitatively analyzing interaction dynamics. We propose four hypothetical types of collaboration that can be characterized by a coarse analysis of the summation dataset, i.e. the creative trajectory monitor curve described above. These categories might be thought of as interaction trends that describe collaborative momentum, since they indicate the general direction of the collaboration in terms of the types of actions being performed up to that point of the interaction. The method for mathematically determining these classifications was inspired by the computational modelling technique used in stock market analysis to determine buy, sell, and hold signals.

7.6.1 Mutual Resource Gathering

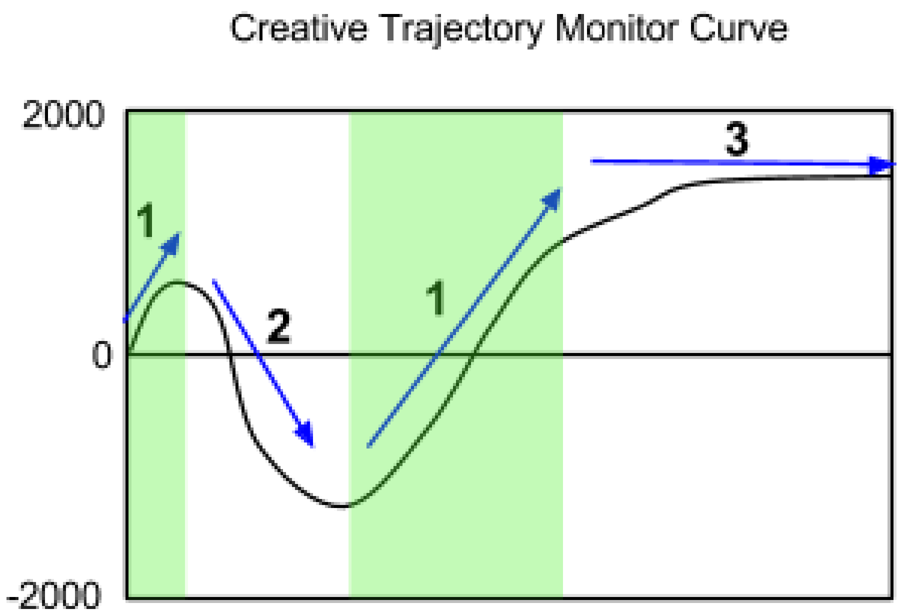


Figure 35: Mutual gathering state where both players are adding new resources.

When both players are unclamped in the positive direction for prolonged period of time (e.g. building structures, gathering resources, discussing the play session), this interaction trend will result. It is is characterized by the creative trajectory curve heading steadily upward (shown in Figure 35). This value can be computationally calculated by gauging whether the fast moving average of the creative trajectory curve moves above its slow moving average. Using stock market modeling convention, this state is classified as a buy signal.

7.6.2 Mutual Waiting/Thinking

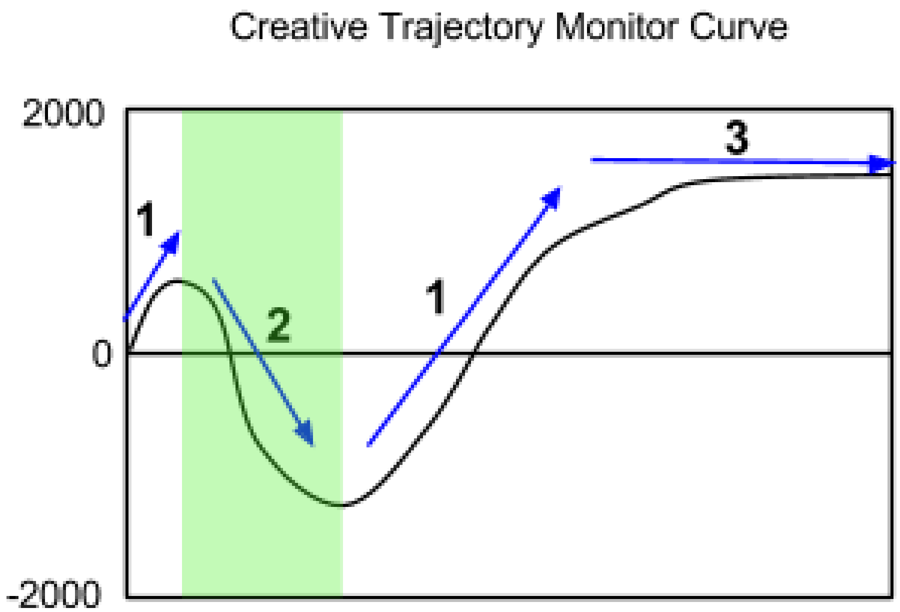


Figure 36: Mutual gathering state where both players are adding new resources.

When both players are unclamped in a negative direction for a prolonged period of time (e.g. waiting or remaining disengaged), this interaction trend will result. It is characterized by the creative trajectory curve heading steadily downward (shown in Figure 36). This value can be computationally calculated by gauging whether the fast moving average of the creative trajectory curve moves below its slow moving average. Using stock market modeling convention, this state is classified as a sell signal.

7.6.3 Quasi-Stationary Play State

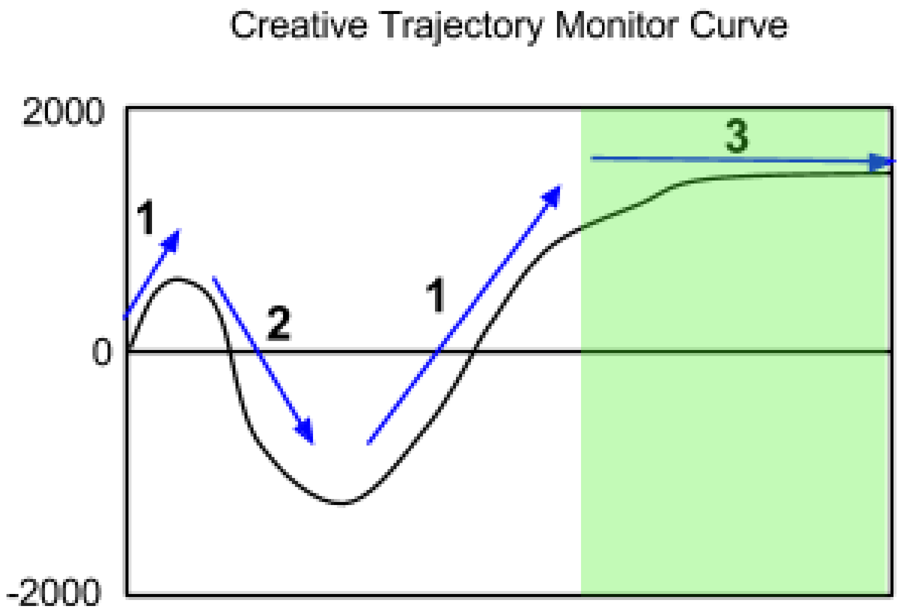


Figure 37: Mutual gathering state where both players are adding new resources.

The quasi-stationary state denotes when the participants' interaction is coupled in some manner. It denotes an time segment of interest for further analsysis. This type of interaction trend is characterized by the creative trajectory curve holding steady by trending horizontally with occasional upward/downward movement within a threshold, possibly with a slight tendency towards a downward trend (shown in Figure 37). This value can be computationally calculated by calculating whether the derivative of the creative trajectory curve is near zero (within a threshold of standard deviation). Using stock market modeling convention, this state is classified as a hold signal. The prediction of a slight downward trend in this state is due to the hypothesis that the ideal coupling between two players is mutual balanced interchange between 0 (play actions) and -.5 (attentively waiting), which would result in a gradual decrease of the creative trajectory through time, until some player restructured the added new meaning to the interaction, such as adding a new character or idea.

7.6.4 Non-Classified

This type of interaction trend is characterized by random deviations, or noise, in the interaction with no clear pattern as determined through optimized thresholds. These segments in time may contain various combinations of 1, 2, and maybe brief periods of 3. The important bounding parameter will be time: a certain period of time must pass in the qualifying condition to be successfully categorized as 1,2,3,4.

This value can be computationally calculated as the remaining segments not classified by the techniques above. The remaining segments will change based on parameters set by the classification algorithms used in the stock market analysis approach as well as temporal thresholds on classifying trends.

7.7 Styles of Coupled Interaction During Participatory Sense-Making

The quasi-stationary play state (interaction trend 3 above) can be described as a type of collaborative *flow* [27]. Many types of interactions could yield this quantitative classification. The next step in this procedure for quantifying interaction dynamics is to tease apart what caused the semi-stationary play state. A few scenarios are immediately obvious and can be broken down into three categories shown below in Figure 38.

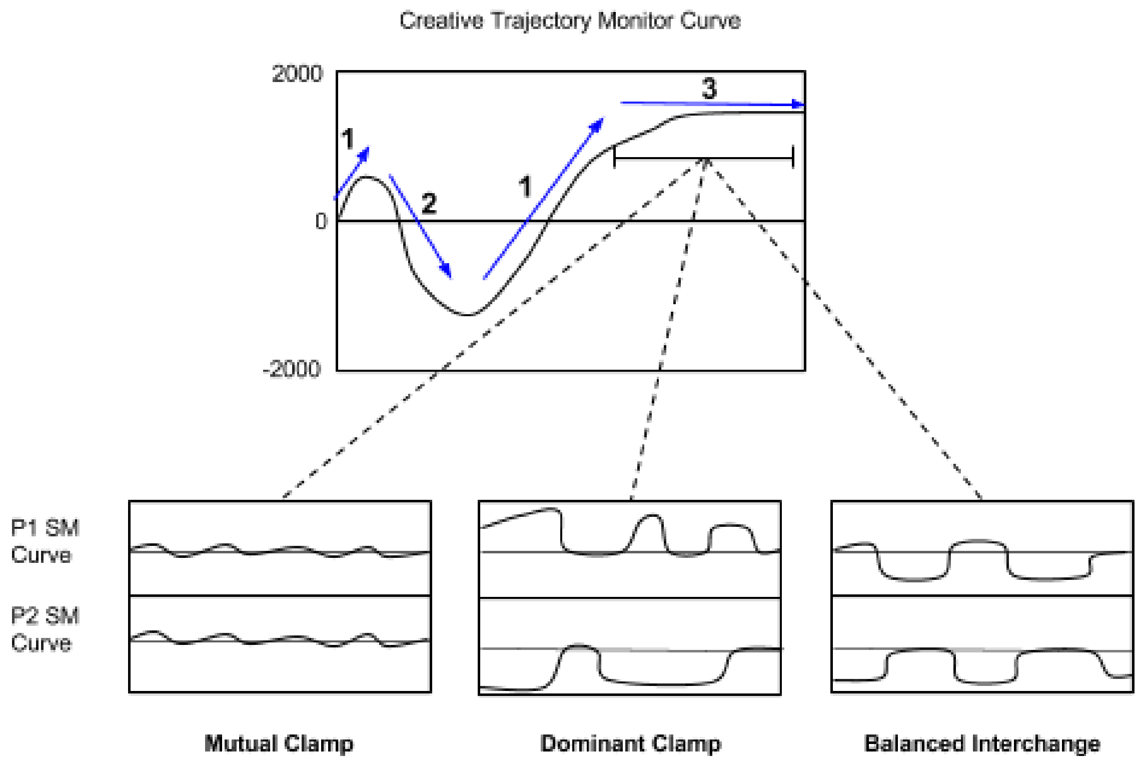


Figure 38: Sub-classifications of coupled interactions that can yield a quasi-stationary state in the creative trajectory curve.

7.7.1 Mutual Clamp

In a mutual clamp scenario both individuals are simultaneously clamped on a play activity/task. This classification can be computationally calculated by comparing the individual value of the sum of the sense-making curves for each player for the segment in question. If the sum of the two sense-making curves for each player during a particular time segment is very similar, then it qualifies as a mutual clamp.

There are two variations within this scenario that require additional analysis. The players could be clamped together or independently. When players are clamped together, they may be working on a joint task or activity. This state can be calculated by referring to observation data (i.e. the video) to determine whether players are working in same area/on same task. Conversely, players may be clamped at the same time working on isolated activities, known as siloed play. This can be calculated by referring to observation data to determine whether players are working in different area/on different task.

7.7.2 Dominant Clamp

In a dominant clamp scenario, one individual is leading the interaction, while the other person is following and not always fully engaged. This state can be computationally calculated by comparing the sums of the sense-making curves. If the total value of one players sense-making curve is significantly larger than the other for the given time period, we can classify this state as having a dominant clamp. Significantly higher summed values signifies that one player was doing all the building and resource gathering for the collaboration, while the other player was passively waiting, or potentially disengaged (while both could still be periodically engaged in play actions).

7.7.3 Balanced Interchange

In a balanced interchange, well formed turn taking emerges whereby each player performs a clamped play action and waits on a response from the other. This state can include other building and restructuring activities in a balanced exchange as well (i.e. one person building, the other watching, and then building something additional).

This type of collaboration is theoretically the most optimal type of exchange in creative collaboration due to its ability to lead to tightly coupled interactions and participatory sense-making with meaning that emerges from both parties. During these couplings, meaning is co-constructed in a way that inspires both parties in unexpected ways, deepening the narrative and immersive potential of the emergent play world and sustaining creative engagement. The coupled interaction yields an emergent autonomy that guides actions and immerses players into an imaginary world with a predictable rhythm that reduces the energy expenditure of the system (by reducing the burden of integrating new information). This type of coupled interaction creates a state of social flow that affords actions through perceptual logic or intelligent perception seemingly automatically and with minimal effort.

This state can be computationally calculated by performing a convolution of the two sense-making curve datasets with one dataset inverted to determine their degree of overlap. Theoretically, the sense-making curve of a perfectly balanced exchange would mirror each other during a given quasi-stationary segment. When inverted, perfectly balanced curves should overlap to a significant degree. These types of balanced interactions would yield a high convolutional value and calculating this value can help indicate the relative *balance* of the interactional exchange. This metric is one potential measure of the quality of a segment of the collaboration. It can be computationally calculated by comparing the overlap between the convoluted sense-making curves. If there is a high degree of overlap, then there was a balanced interchange.

7.8 *Exemplar Dataset and Analysis*

This section shows the application of the sense-making curve analysis technique on a real data point from the pretend play session with 32 adult dyads. This data represents the sense-making curves coded for each of the individuals in the play session.

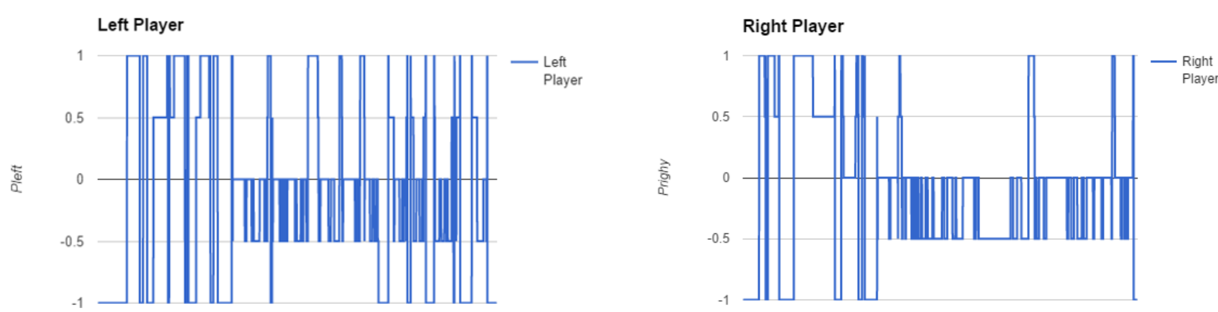


Figure 39: Sense-making curves from the right and left player of a play session

The sense-making curves shown in Figure 39 show the qualatative coding classification of two participants during a 5 minute play session. Here, we can see both players greatly fluctuating above and below the axis (i.e. building meaning/restructure environment and waiting/disengaged) during the first third of the play session, which most likely correlates to their setup period during which they are figuring out what kind of activities to engage in during the play session by exploring the playmat, toys, and resources at their disposal. This early period represents more active sense-making where the participants are actively experimenting with the environment and directly communicating about what to do (i.e. extradiagetic communication). After about 1:30 (about a third of the way through the x-axis), both participants begin to engage in play activities and their curves fluctuate more closely around the 0 axis (i.e. clamped cognition), cycling between 0 (engaged in fluid play activity) and -.5 (attentively waiting). This trend indicates some degree of turn-taking is beginning to emerge at this point. It is important to note that the left player continues to deviate greatly from the 0-axis at distinct points in both the positive and negative directions, suggesting that this player was more actively involved in adding new elements into the play session as well as disengaging more often, especially in the last third of the play session.

The next step in the analysis is to take a cumulative integral of each of these sense-making curves to provide a more global picture of the participants' activities through time. For the purpose of this exemplar demonstration, the Matlab cumsum function was utilized to approximate the integral. In more precise analysis, the Matlab function cumtrapz is used to increase the accuracy of the integral approximation.

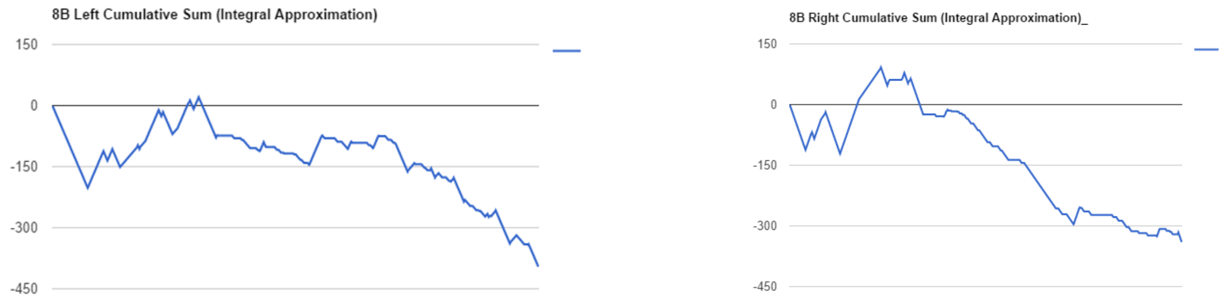


Figure 40: The cumulative integrals of each player in a play session.

The cumulative integral approximations shown in Figure 40 provides an easier way to visaulize how each participant was contributing to the overall flow of the play session activities. From this representation, it becomes clear that both players experienced a relatively dramatic initial hesitation and waiting period followed by a period of building meaning in the environment. However, after this initial phase, their actions seem to diverge as the right player spends more time watching or hesitating, while their partner, the left player, continues to build new meaning and engage in play (as represented by periods of rising during the culuative integral and holding stead, respectively).

The next step in the analysis is to combine these two cumulative integrals together to provide information about how both of the participants actions related to each other through times. To accomplish this operation, two one-dimensional arrays (or vectors in Matlab terminology) are added together. This combined cumulative integral has the capacity to begin to identify some interaction trends involving both players.

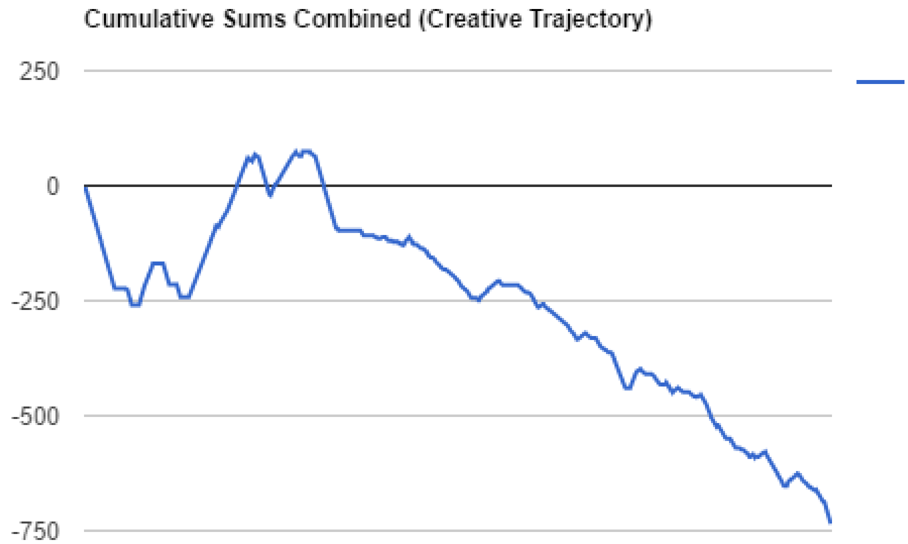


Figure 41: A graph depicting the combination of both participants cumulative integral curves.

The graph in Figure 41 shows the combined cumulative integral of the participants' sense-making curves, which shows the overall activity through time with respect to both player's actions. We refer to this curve as the creative trajectory because it can be used to help identify interaction trends and patterns in the overall creative flow, or trajectory, of the play session.

The next step in the analysis is to identify the four main interaction trends in this creative trajectory curve. For the purpose of this examplar demonstration, these trends are visually identified and depicted on the graph. In the actual analysis, more precise computational modeling techniques are used, similar to the conventions used in stock market analysis to identify buy, sell, and hold signals from continuous stock data.

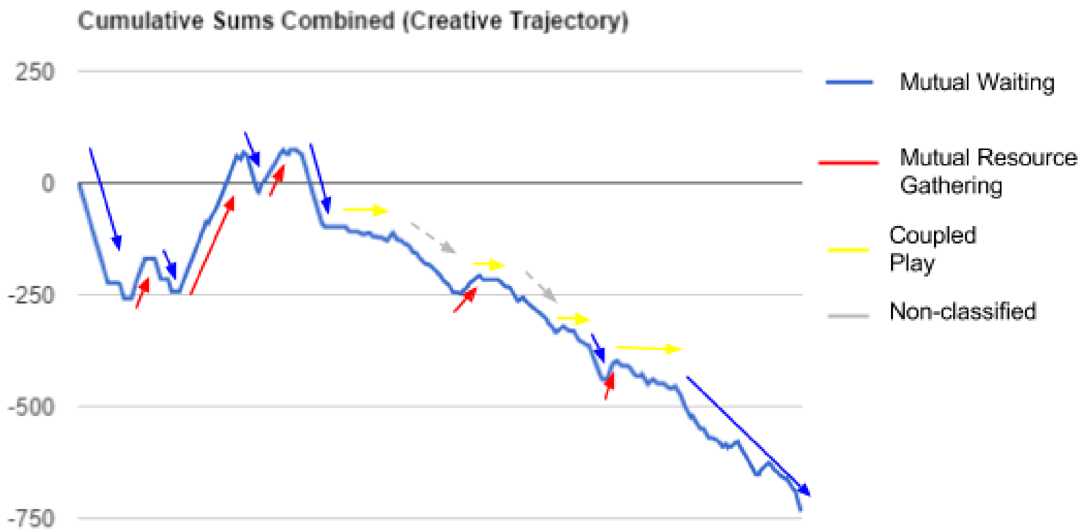


Figure 42: Interaction trends that can be classified by analyzing the creative trajectory curve.

The graph in Figure 42 shows how different interaction trends can be identified using the creative trajectory curve. Each of these trends theoretically corresponds to different ways of collaborating as well as helping to identify when coupled play began to arise between the players (the yellow arrows in the graph). From here, many additional analyses can help describe the interaction dynamics of the play session. The trends visualized on this graph were explained earlier and can be summarized as follows:

- **Blue Line: Both players are hesitating or waiting.** This classification will be most distinct when both players are completely disengeaged for prolonged periods of time as they will both have values of -1 during those times. When both players are coded as -1, the cumulative integral is reduced sharply, as indicated by the large negative slop in the first three blue arros. However, the value of the cumulative integral can also fall when both players are either attentively waiting (coded as -.5), or one is attentively waiting while the other plays (coded as 0). In this case, the slope of the creative trajectory will not fall as fast, such as the last blue line on the right side of Figure 42.

- **Red Line: Both players are building meaning or gathering resources.** This classification coresponds to both players actively engaging in a sense-making process (unclamped cognition) through interactively exploring their environment by building new structures in the play area or looking for new resources to add to the play area. When the slope of the creative trajectory is positive, both players are most likely engaged in unclamped actions (rated as +.5 or +1) to make the cumulative integral rise. A high positive slope would indicate that both players are actively looking for new resources (i.e. rummaging through the toy box) to add to the play mat. A smaller positive slope may indicate building or tweaking elements that are currently on the play mat.

- **Yellow Line: Clamped quasi-stationary play state.** This classification represents some type of coupled play interaction between the participants. There are many different types of coupling that can result in this type of semi-stationary state where the slope of the creative trajectory curve is near zero. To classify this state, the derivative of the creative trajectory curve can be calculated and those areas of the derivative curve that are near zero within a threshold of standard devisation can be classified as this state. These couplings indicate regions of interest upon which to perform further sub-classification analyses.

- **Gray line: Non-classified states**. This classification represents those points of the play session that do not fall under the three classifications made above. These ambiguous states of play may fall into one of the above categories when the tuning parameters for the classification algorithms are changed. They do not have to be directly considered in the analysis, but their presence can help researchers understand the granularity of their classification parameters.

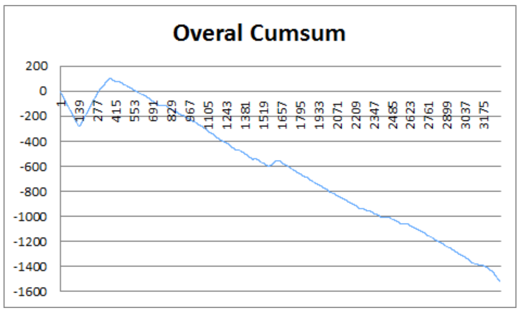
The next step in the analysis is to identify which periods of time participants were engaged in coupled play (yellow classification line) and use those time segments to go back to the initial sense-making curves to analyze the original sense-making curve values to see what types of activities participants were engaged in during the coupled interaction. For example, were both players simultaneously clamped on play actions? If so, were they mutually clamped on the same activity, such as attacking a castle? Another explanation could be one person was continuously adding meaning while the other person was continuosly in the attentive waiting stage, which would result in the cumulative integral sum remaining largely the same and yielding a quasi-stationary state in the creative trajectory curve. The values from the original sense-making curve will reveal these nuanced details, allowing further sub-classification of the quasi-stationary play state into: dominant clamp (one person is adding new meaning and directing the play), mutual clamp (both participant are playing simultaneously), and rhythmic balanced interchange (participants are taking turns).

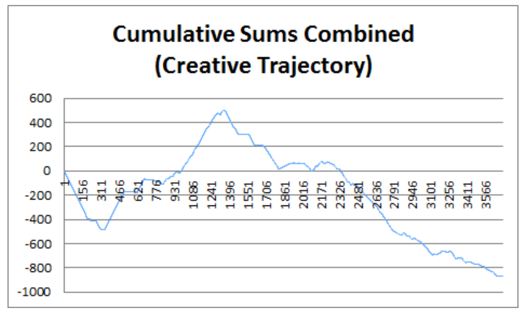
From this point in the analysis, it is also possible to count how much time was spent in each type of interaction trend and when each interaction trend tended to occur within the overall timeline of each pretend play session. With this data, it may be possible to identify more or less successful collaboration strategies by comparing these broad interaction trends with the initial 'play score' each video was given during early stages of analysis.

7.9 *Initial Pretend Play Study Interaction Dynamics Analysis*

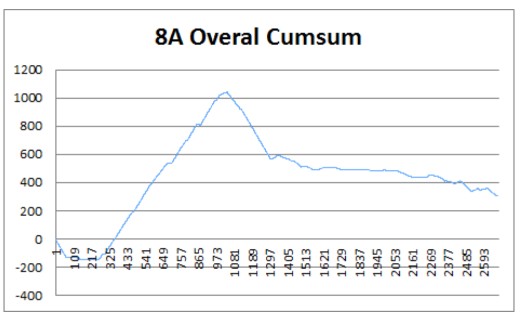
The next dataset shows the creative trajectories from the five highest and lowest rated videos from the adult dyad pretend play study. A visual inspection provides a better idea about how this data analysis method can be used on larger datasets.

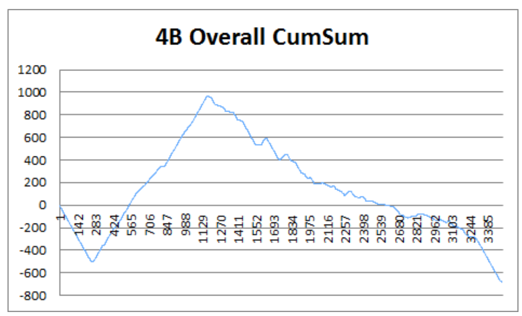
Highest Rated Play Sessions	Lowest Rated Play Sessions
Session 3B	Session 4A



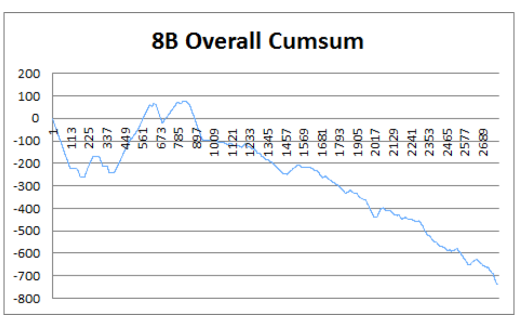


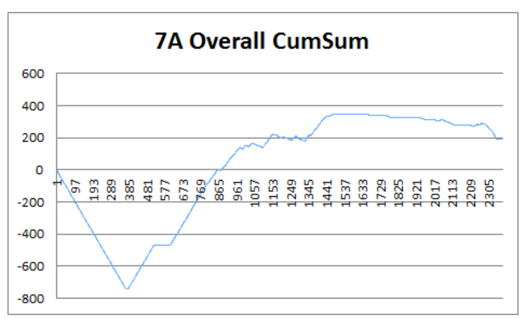
Session 8A	Session 4B



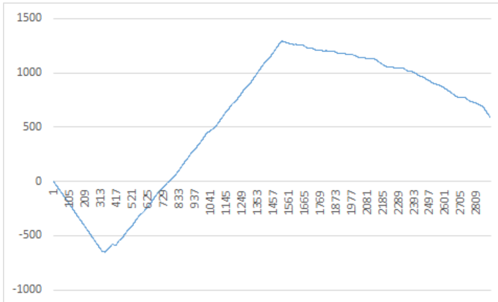


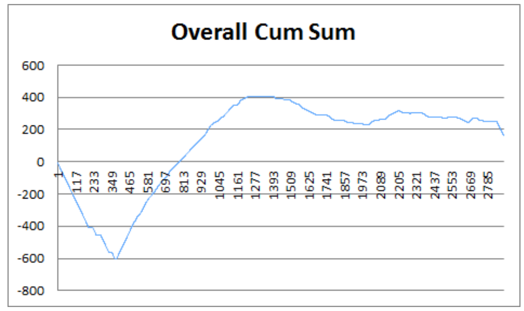
Session 8B	Session 7A





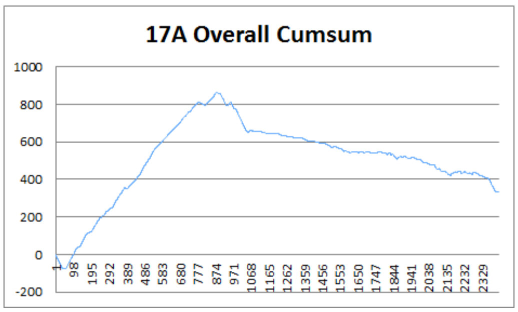
Session 10B	Session 11B





Session 17A

Session 12A



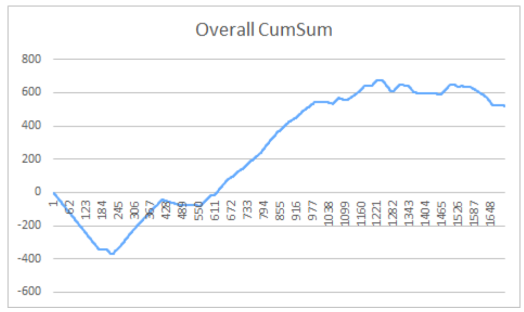


Table 8: Comparison of the creative trajectory curves from the five highest and lowest rated play sessions.

Table 8 shows the creative trajectory curves from the top five highest and lowest rated play sessions as determined by subjectively viewing the sessions, gauging their overall success, and comparing those rating between codings to resolve discrepancies (prior to and independent from coding the sense-making curves for each session). Given the variety of interaction styles and collaboration techniques that can arise during open-ended improvisation, we do not expect to find a single style of collaboration or interaction pattern that is responsible for success. Instead, we hope this technique can help quantify different sense-making patterns and help understand the overall dynamics of the collaboration, whether successful or not. From this subset of curves, we can begin to identify some critical features that are common between the curves in both successful and unsuccessful play sessions.

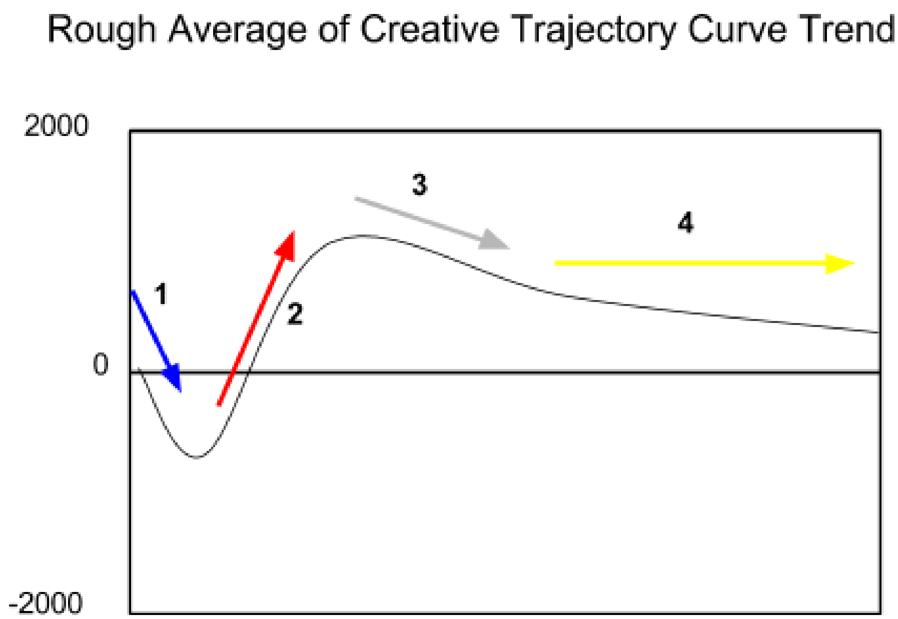


Figure 43: Approximate average creative trajectory based on both high rated and low rated videos

From this data sample, there appears to be four main trends that occur in a relatively well established order. In all of the datasets sampled through this ranking procedure, the interaction begins with a downward trend creating an initial dip (as labeled 1 on Figure 43). Following this dip is a general increase in value (shown in 2 in Figure 43). Then, there is a peak and reverse in slope (shown in 3 in Figure 43), followed by some sort of stabilization or general decrease (shown in 4 in Figure 43). There are deviations from this general trajectory, and it is interesting to investigate whether there are marked differences in this general curve between the successful and unsuccessful play sessions. These phases can be directly mapped to the sense-making processes we identified in our enactive characterization of pretend play remunerated below:

1. **Prepare the mind:** before participants begin to play, they first orient themselves to the environment and generally prepare mentally to figure out some

kind of strategy or approach for interacting in the current situation. This is generally characterized by participants hesitating and not quite knowing how to act early on in the play session, perhaps with awkward pauses.The magnitude and length of this period may play a role in the quality of the collaboration. Given the evidence above, it appears there is at least a weak correlation between the size of this phase and the outcome of the play session (as the low ranked play sessions generally had a larger time spent in this first phase). For example, it may be the case that those players that will not be very engaged or committed to the play session will spend a longer time in this preparation period.

2. **Build meaning:** after the preparation phase, players begin to actively explore and transform the environment by gathering resources (toys, blocks, etc.) and building meaning (physically and verbally) into the environment through interactive experimentation and extra-diegetic communication with their partner. The magnitude of this meaning building phase seems to change drastically between different participants. For example, in Table 8, in session 17A this phase is practically non-existent, but in 10B, this upward trend comprised the majority of the session. Both of these play session were rated as highly successful, which highlights the variety of collaboration styles that acn yield successful play.

3. **Enact the narrative:** The building phase described in (2) is typically sharply punctuated by a change in slope of the creative trajectory from positive to negative. This generally corresponds to when participants begin to enact a narrative (based on foundation of nucleus meaning built in the preceding phase) by taking on the persona of the characters involved to act out elements of an unfolding story. At this point in the overall dataset, the trends seem to diverge more dramatically, sometimes exhibiting coupling and at other times exhibiting

gradual and sharp increases/decreases. Sharp increases would theoretically represent a point where new meaning was introduced into the environment, which may provide some indication about the diversity of ideas explored during a play session.

4. **Maintain the flow and deepen the narrative:** A good collaboration requires maintenance throughout the session. Depending on how the course of the interaction has unfolded until this point, individuals sometimes introduce new ideas and think up alternative approaches to help keep the play session interesting. It is not immediately clear without a more systematic analysis (which is currently underway) what constitutes a good or bad way of maintaining the flow of collaboration as this is a very complex question. However, one thing that can be immediately noted by inspecting the dataset is that some curves obviously exhibit more continuous trends whereas others have more noise and variance. This can be partially due to the limitation of the coding technique (e.g. differences in the granularity with which the initial codes were applied to the video), but it may also be linked to prolonged interaction trends that may help understand different strategies and approaches for effectively maintaining the flow of an open-ended improvisation. Finding periods of increased slope followed by a quasi-stationary play state, for example, may indicate a deepening event where a new piece of meaning has been successfully introduced to the emerging narrative.

Graphically representing the creative trajectory curves and analyzing their trends can serve as a powerful data visualization and conceptual aid to reason about open-ended collaboration improvisation. While these general trends provide novel insight into the nature of interaction dynamics in open-ended collaboative improvisation, the continuous nature of the sense-making curve data provides opportunities to explore much more granular analysis and classification techniques. Further classification using the computational modeling convention described in this chapter (i.e. the stock market analysis technique) allows further quantification to facilitate comparisons throughout the dataset. The present analysis demonstration and discussion should serve the purposes of this chapter, which was laying out the need, theoretical justification, coding method, and initial application of the creative sense-making analysis technique.

7.10 Initial Drawing Apprentice Study Interaction Dynamics Analysis

This section reports the preliminary results of applying the sense-making curve technique described in Chapter 5 to the data collected during the creativity study described earlier in this chapter. This analysis helps quantify the interaction dynamics of the user and the co-creative agent during the study. This analysis was conducted about a year after the initial data collection from the Drawing Apprentice creativity study. One reason for conducting this analysis was that it was difficult to tease apart exactly how each partner (both agent and wizard) affected the quality of the collaboration. Event-based coding procedures provide information about the number of times events occured, but it is hard to draw conclusions about what feature of the interaction were responsible for the observed events. Since collaborations unfold through time in an open-ended way, activities and ways of interpeting them dynamically change through time–the history of the interaction can play a crucial role in the behavior later observed. However, without a continusous source of data describing the user and agent's interaction, it is very difficult to account for these temporal features of interaction dynamics.

Using the sense-making curve coding technique allows researchers to collect and analyze continuous data describing interaction dynamics and cognition, which facilitates the use of continuous mathematical functions in their analysis. The continuous nature of the data helps account for the temporal dimension of open-ended collaboration, providing information about the rhythm of interaction, style of turn taking, and strategy for mutually making sense of the situation.

We first present the full sense-making analysis of one participant from both collaboration conditions (i.e. agent collaboration and wizard of oz collaboration). Then, we present the creative trajectory curve for all wizard collaborations next to their corresponding agent collaboration to observe differences between the interaction dynamics in both conditions.

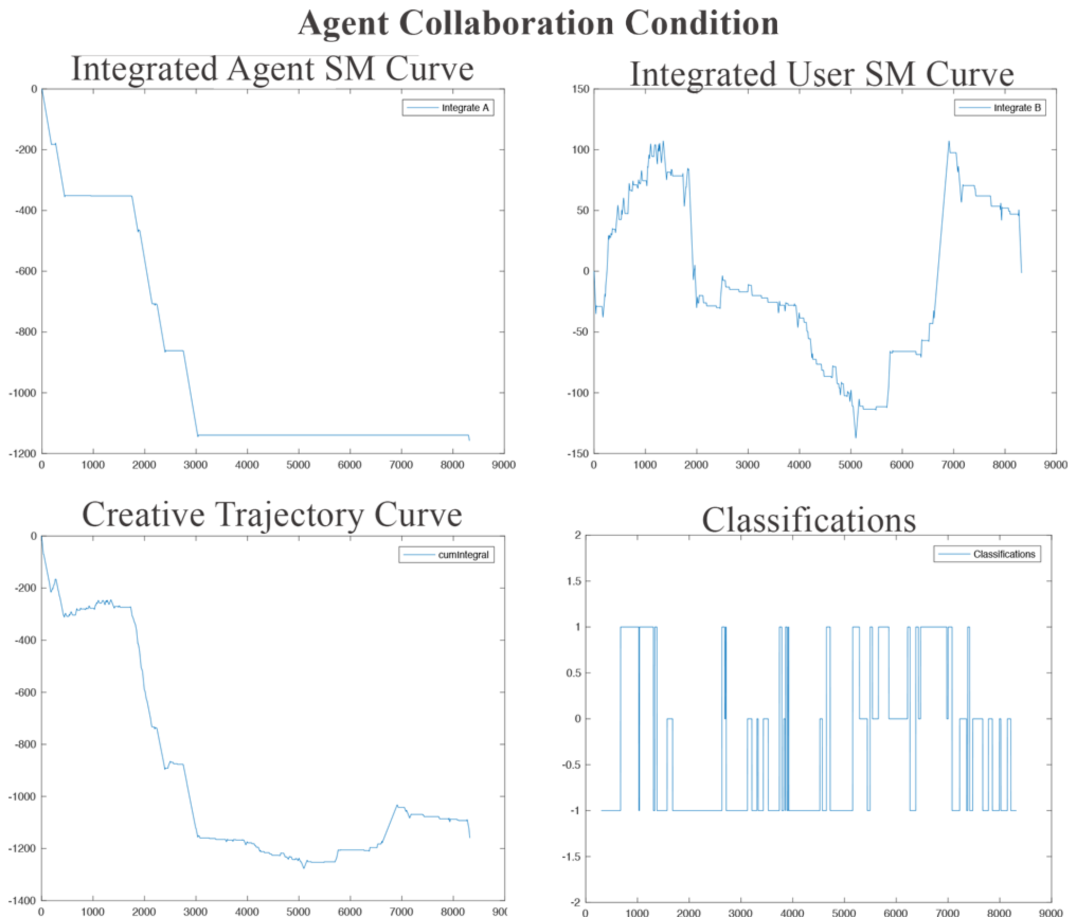


Figure 44: Interaction Dynamic Data from Participant 1 Session 1: Collaboration with the Co-Creative Agent

The interaction dynamic data from the agent collaboration condition of participant 1 shown in Figure 44 has some interesting characteristics and is quite atypical among the data points collected from the study. For example, the integrated sense-making curve of the agent (top-left of Figure 44) shows three distinct sharp downtrends, which signify the agent was inactive for extended periods of time. In the program, this occurs when the agent is waiting for the participant to finish long turns. Additionally, there are three distinct horizontal periods, with an extremely long segment in the second half of the drawing (approximately 6 minutes). These periods all represent continuous drawing from the agent. That last segment represent's the agent's long turn in response to the user's many line inputs while it was waiting.

The integration of the user's sense-making cure is equally revealing (shown in top-right of Figure 44). There are two large unexpected spikes in the data. Small increases in the user's integrated SM curve normally represent short interactions with the interface, such as providing feedback, and changing colors, thickness, and drawing modes. However, this participant's integration has sharp and large increases in value. After looking back to the video data to investigate this matter further, it was discovered that the participant was watching the agent during it's particularly long turn and experimenting with changing the colors the agent was drawing with as it was drawing. This was not an intentionally designed feature of the tool, but this participant seemed to enjoy having active control over what the agent was doing while it was drawing.

The retrospective protocol interview data for this participant revealed that the user was frustrated with the system during this part of the session. She wanted it to stop drawing and end its turn, but she couldn't figure out how to get the agent to stop drawing. (There was actually no way to stop the agent in the program. Since then, we have experimented with explicit stop and start mechanisms with mixed results as starting/stopping turns can be tedious.) In this case, the participant voted down thinking the negative feedback would end the system's turn, but became more frustrated when the agent seemed to not pay attention to her feedback. However, once she began adjusting the colors of the agent, she became more engaged in that activity for a while, accounting for the second large spike in the session. Eventually, the system ended it's turn and the curve shows more contributions from the participant as horizontal lines.

The creative trajectory curve represents the sum of the two integrated sense-making curves shown in the figure. However, the vertical scales of the two integration SM curves are different. The plot in the top-left has a vertical range of 1200, while the plot in the top-right has a range of 300. A shared axis helps compare the same curves side-by-side, but some of the detail is also lost from that graphing convention. The current graphs use the default plotting convention of Matlab that select axis size based on the range of the values to be plotted.

The classifications plot in Figure 44 applied the stock market computational modeling convention described in Chapter 5 to the creative trajectory curve pictured. In the classification plot, 1 represents the mutual gathering upward trend, 0 represents a quasi-stationary coupling, and -1 represents the mutual hesitation downward trend. As noted earlier, there are more than expected instances of the increasing trend (1) in the dataset pictured. Originally the upward trend classification (1) was meant to denote when both players have positive values of their sense-making curves. However, with the Drawing Apprentice, the agent can never actually achieve a 1 score since it does not actively experiment with it's environment (in a perceptible way) or communicate to the user about what it is doing (i.e. explainable AI). Therefore, in the current dataset, the co-creative agent can only achieve ratings of 0 (actively drawing) and -1 (waiting) when using the human behavioral marker code application metrics established by the clamping/unclamping convention.

As a result of this coding convention, we expected there would not be many general upward trends in the creative trajectory of the user and agent since the agent was not contributing any positive values to the cumulative sum. That prediction holds true on average for the dataset, but there are a few instances that deviate that prompted additional investigation.

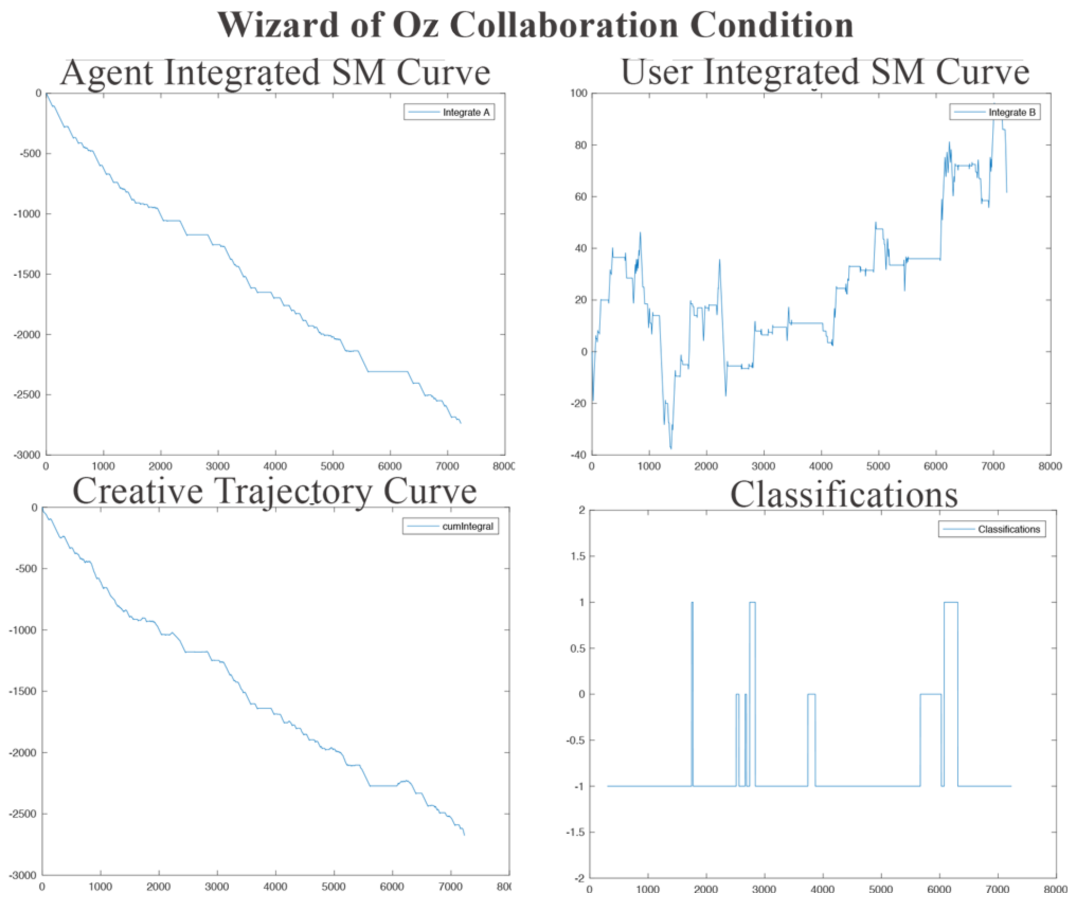


Figure 45: Interaction Dynamic Data from Participant 1 Session 2 Collaboration with the Wizard

In the wizard of oz study data from participant 1 shown in Figure 45, I was controlling the agent (as an expert collaborator). The goal was to mimic the style of the agent, such as its turn taking procedure and guidance from feedback, while being able to understand what users were drawing and respond intelligently. This approach was selected instead of a freeform collaboration to control the conditions of the two scenarios as close as possible. The user was informed they were using two versions of the same system with different drawing algorithms. The intention behind this approach was valid, i.e. wanting to study whether the system was able to collaborate in a similar was as a human. However, when restricting the artist's decision as the 'wizard' (to keep conditions similar), the artist's natural collaboration style is significantly reduced, drastically reducing the potential diversity of data and quality of collaboration, as evidenced by the consistent downward trend of the creative trajectory curve shown in Figure 45 and seen throughout the data (shown in Figure 46). When conducting wizard of oz style studies in the future, a three condition approach may be more suitable: 1) User collaborates with system; 2) User collaborates with 'wizard' through the system; and 3) User collaborates face-to-face with the artist that served as the wizard in the experiment using pen and paper.

The user's integrated curve in Figure 45 show some sharp increases, but the range is only 100 for the plot. As a result, the deviations shown in the user's SM curve are not as impactful as the agent's SM curve in the final creative trjaectory curve shown in the bottom-left. Overall, there were many downward trends (-1) classified in this dataset. It is important to note the these results, especially the classification results, are still preliminary. We have not yet determined the optimimal parameters to use in the stock market classification scheme. However, the approach we present appears to show promise as it successfully identifies significant deviations from a norm, such as the case of participant 1 becoming frustrating and beginning to play with the interface controls while the agent was drawing.

When examining Figure 46, it is immediately obvious that P1 had a significantly different collaboation experience than the other participants. The slope of both the agent and user's cumulative integral is relatively flat compared to the consistent downward slope of the other participants. Between the agent and wizard condition in P1, we see that the wizard did not have extremely long turns like the agent, but the user still had long turns (flat sections of the curve) and exploratory behavior (increases in the curve). In fact, comparing the wizard cumulative integral curves to the agent cumulative integral curves overall, we can see that the wizard does not deviate from a consistent downard slope, while the agent does in P1 S1, P4 S2, and slightly in P5

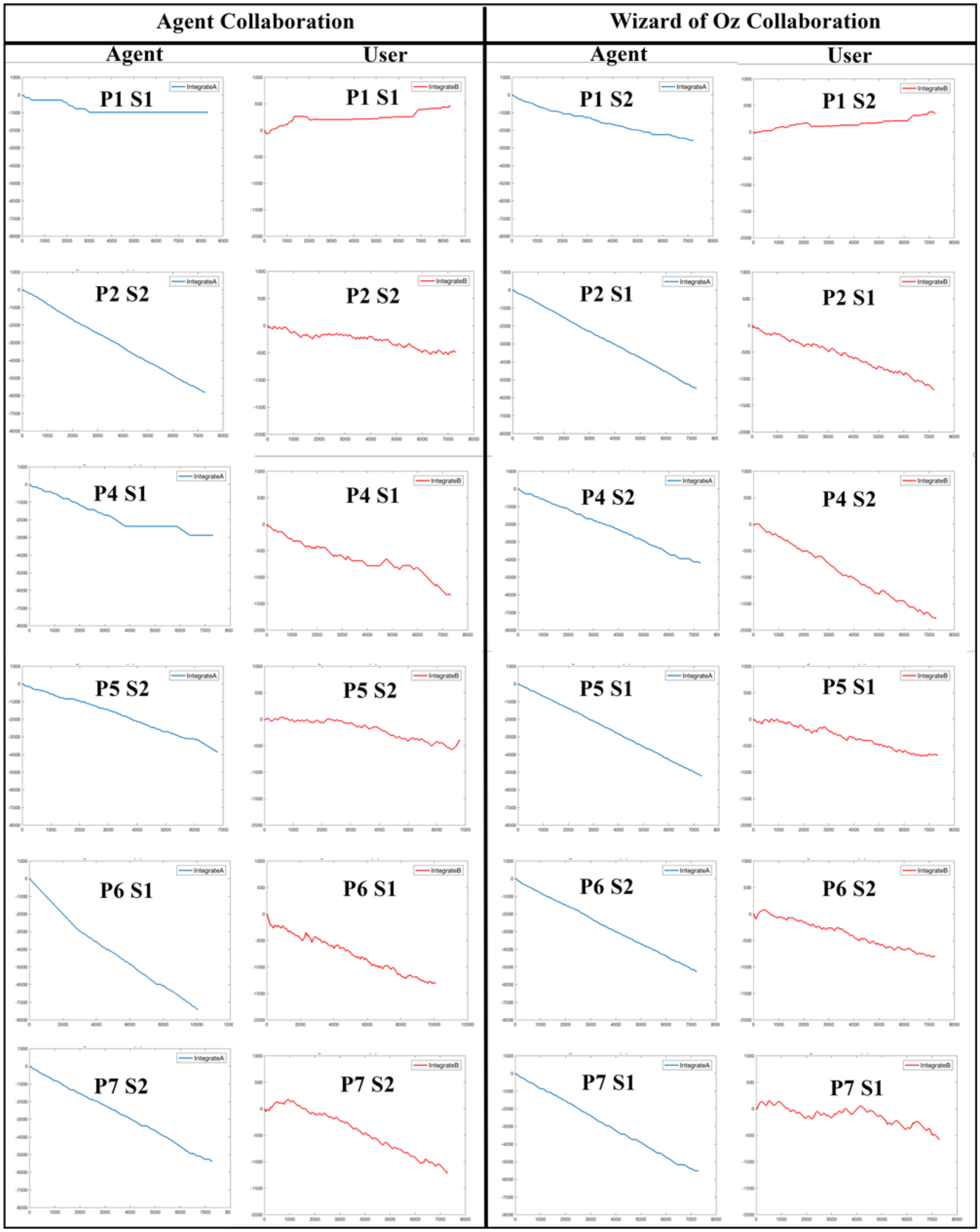


Figure 46: The cumulative integral of the agent and user's sense-making curves from six user studies in the Wizard of Oz and agent collaboration conditions.

S2. The wizard is more consistent in its curve because it is not responding to literally every single line of the user like the agent is. This feature of the agent's behavior can result in very long turns, exemplified by the flat sections of the SM curves mentioned previously.

It is also apparent in Figure 46 that the agent (both wizard and agent conditions) has a larger negative slope than the user overall. This is due to the fact that participants were typically in an 'attentively waiting' state while the agent was drawing, but the agent is always in a 'completely disengaged' state when the user is drawing. Small interface designs that indicate the agent is waiting (such as a timer) would alleviate this difference.

Comparing the cumulative integral chart in Figure 46 to the next Figures describing the overall creative trajectory for these sessions (Figures 47 and 48 also provides some insight. First, it is apparent that the classifications outlined in Chapter 5 will not be that productive when applied to the creative trajectories form this user study since there is minimal variation in the downard trend of the creative trajectories. However, there are significant variances in each of the individual cumulative integral curves that can be classified using the same technique outlined in chapter 5 for the overall creative trajectory. Increasing the agent's ability to express its cognitive state would provide more variability in its curve and make the creative trajectory appear more like human collaboration.

Examining Figure 46 with the stock market analysis technique in mind, we see several instances of quasi-stationary segments in P2 S2, P4 S1, P5 S2, and P7 S1. If the agent had more degrees of expressive freedom (rather than completely disengaging between turns), these segments could have resulted in a detectable mutually co-regulated coupling. However, we can still gain insight from the cumulative integral curves paired with qualitative insight. P2 is a particularly interesting case since the participant tended to employ one line turns throughout both conditions. This participant reported being heavily influenced by the agent's decisions (in both conditions) and appeared to build off the agent's contributions significantly based on her retrospective reflection. P5 was also an interesting case in that this participant established 'perceptual logic' with the agent by doing certain types of activities in particular regions (e.g. wavy lines, drawing circles, and filling containers with circles). The qualitative analysis showed that the participant was engaged in a type of mutually co-regulated coupling during these instances, and the amount of quasi-stationary segments in P5's curve support this qualitative report.

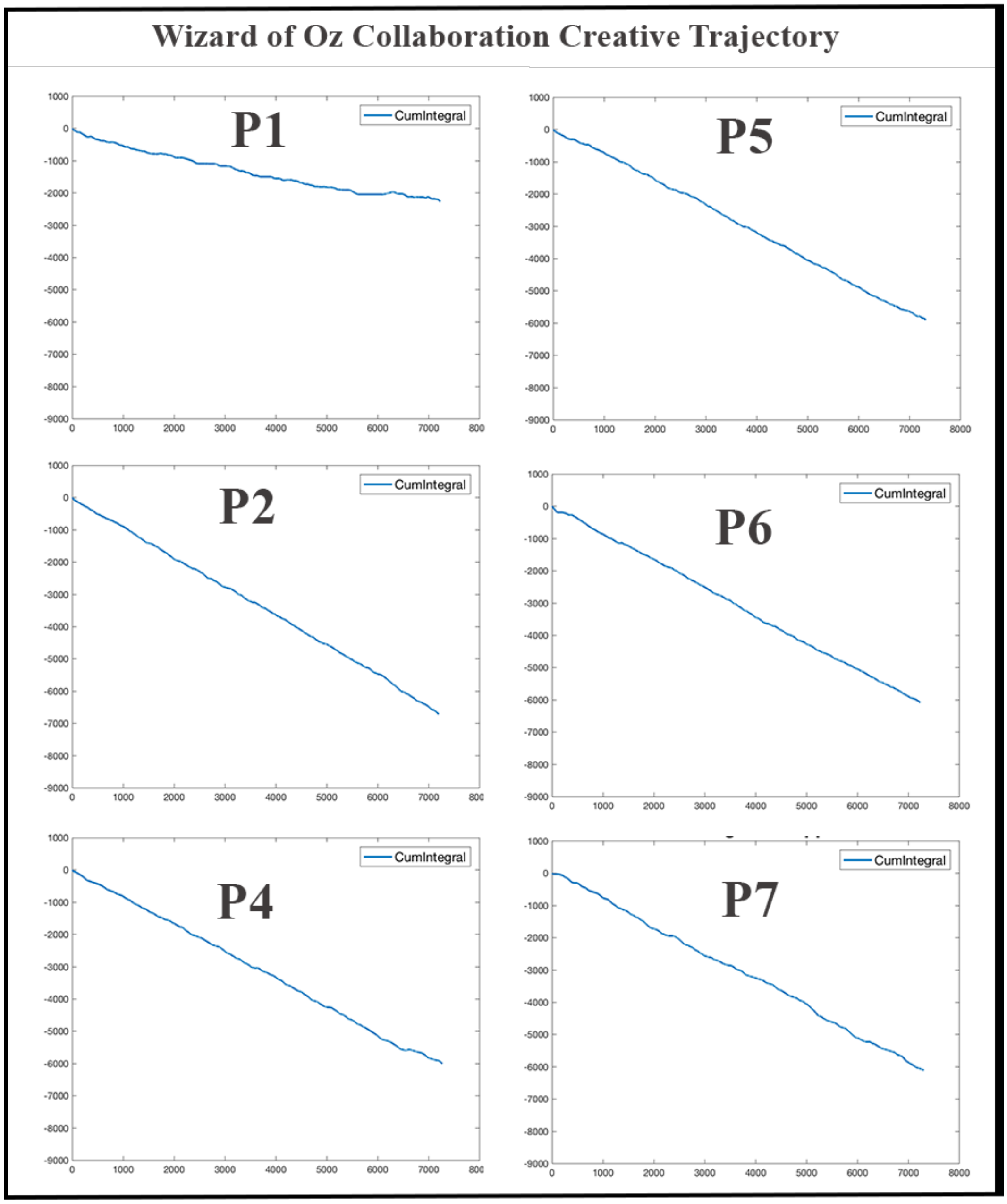


Figure 47: Creative trajectory curves (summed cumulative integral of both players' sense-making curves) from all six sessions in the Wizard of Oz collaboration condition.

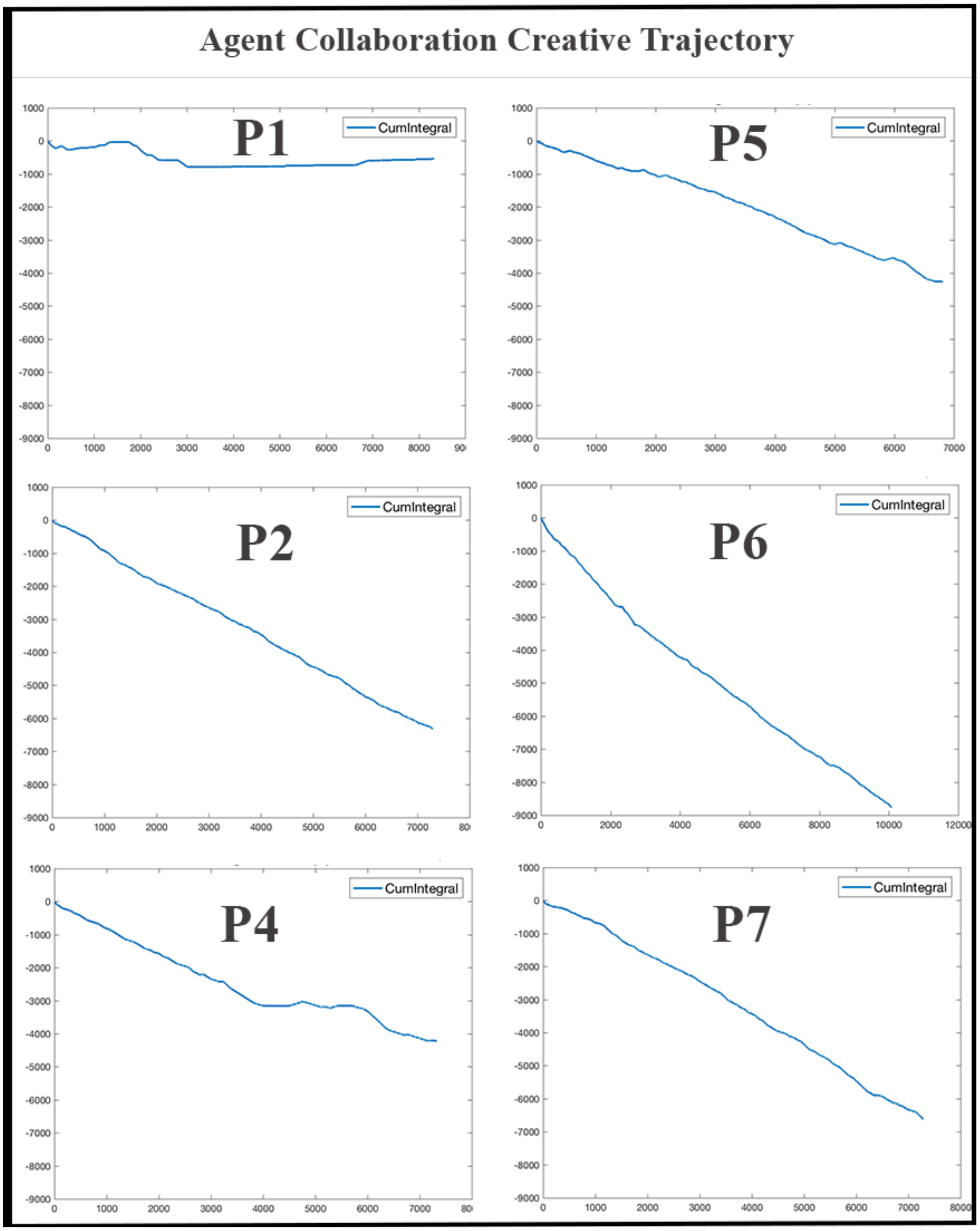


Figure 48: Creative trajectory curves (summed cumulative integral of both players' sense-making curves) from all six sessions in the agent collaboration condition.

In Figures 47 and 48, the slopes of most of the creative trajectories are downard, seemingly at a similar rate. P1 and P4 are obvious outliers in the data due to their deviation from this overall trend. The most obvious question to ask is: why are the slopes generally heading in a negative direction? One large reason is because the system is continuously waiting for the user to act without making any expressions or indications as to its intentions, i.e. it is disengaged. While it is waiting, it is continuously generating a -1 value when it is coded by the analyst. These values significantly impact the cumulative sum score between the user's SM curve and the agent's SM Curve. During a human drawing collaboration, the creative trajectory is predicted to look more like those from the pretend play study presented in Chapter 5. The ability for humans to interact to build meaning through communication and bodily feedback adds a critical ingredient to naturalistic collaboration.

This finding aligns with qualitative reports about users wanting a timer to know when the systerm is going to take its turn and possibly some way to stop it while it's acting. A timer would serve as a means of communication to indicate the system is *attentively waiting*, which would earn it a -.5 score according to the behavioral coding scheme. Design recommendations stemming from this downtrend generally relate to enabling the system to communicate its mental state and reasoning process. Another way to achieve this would be to create a simulated perceptual field for the agent to inform the user what the agent is paying attention to, and to allow the user to direct the agents attention (and subseqently its behavior) to different regions.

7.11 Discussion

To make the creative trajectory curves from the Drawing Apprentice system we investigated appear more naturalistic, the agent needs to have more capabilities to express its cognitive state (e.g. visibly waiting for the user, communicating its interpretation of the drawing, communicating its intention, etc.). We also advocate a new WoZ user study design that includes a few new conditions, such as introducing a baseline and making the wizard condition more naturalistic, meaning that wizards collaboration decisions are unconstrained by how the agent works. Future evaluations of the Drawing Apprentice system (and other similar co-creative systems) could employ the following experimental design for WoZ studies:

- **Baseline Condition:** Individual free draw task to provide a baseline to quantify the interaction dynamics of the participant's creative process without any intervention
- **Agent Collaboration Condition:** Collaborate with the co-creative agent
- **Wizard Collaboration Condition:** Collaborate with a wizard (expert artist) that is controlling the response of the system. Importantly, the wizard should employ their naturalistic collaboration techniques rather than trying to fit their artistic and collaboration style to the system. For example, if artists typically begin their collaborations with 1-line turns and gradually move to synchronous drawing, they should feel free to employ that strategy with the participant, even if the system does not employ that type of procedure.
- **Artist Collaboration Condition:** Collaborate with the artist that was controlling the wizard face to face with pen and paper. Collaborating through an interface drastically reduces the subtle feedback and communication that is mutually available in a face-to-face collaboration. Including this condition provides data about the most naturalistic interaction dynamics of collaborative drawing, resulting in a proper ceiling with which the agent and wizard conditions can be evaluated against.

These new conditions would provide a spectrum of expressivity that would result in several different types of sense-making curves to compare these conditions in depth. Face-to-face collaboration allows the full range of human feedback (e.g. language, gesture, facial expressions, paper movement, etc.). Naturalistic collaboration with a wizard removes all human feedback but still allows the expert artist to collaborate in a flexible and naturalistic way without any channels of communication. The agent condition would remain the same. Introducing a baseline condition would help understand how the participant tends to engage in drawing activities without any intervention. With this experimental setup, there would be multiple levels of feedback fidelity that produce progressively less rich (i.e. varied) sense-making curves, and we predict this experimental design would improve the analytic power of such a study in the future.

After each condition, participants should be asked to engage in a retrospective protocol analysis during which they watch a video of their collaboration and provide commentary about their experience. During this retrospective protocol analysis, it may be possible for researchers to create a rough sense-making curve with the participant. The researcher could briefly explain the conventions behind the sense-making curve and ask the participant to provide commentary about their mental state with respect to the sense-making curve. The sense-making curve may serve as a tool for mutually understanding and communication about the complex cognitive processes occurring during the creative process.

7.12 Conclusions

This chapter presented the technical need for the proposed sense-making curve coding and analysis technqiue. The theoretical justification for the choice of the conventions for behavioral coding markers as well as graphical representation were described. We showed how the continuous nature of the sense-making curve data allows continuous fucntions, such as integrations, to provide information about how the overall play session is unfolding through time. This technique was applied to an exemplar data set to demonstrate the type of data that such an analysis yields. Next, the creative trajectory curves of the highest and lowest rated play sessions were presented and compared. A general trend was put forward from this sample of the data that aligns with the stages of the enactive characterization of pretend play presented in Chapter 4. The sense-making curve technique will be used later in Chapter 7 to analyze user study data from the Drawing Apprentice studies.

CHAPTER VIII

FUTURE WORK

8.1 Summary

This chapter summarizes the future directions for the work described in this dissertation. These research trajectories are organized around the three main thrusts of this work, namely applying the theoretical framework of participatory sense-making to open-ended creativity, quantifying interaction dynamics, and exploring the interaction dynamics and technical approaches to facilitate participatory sense-making in co-creative systems. The chapter begins by describing extensions of the Drawing Apprentice systems based on empirical investigations and user needs. Next, it describes how the creative sense-making cognitive framework can be further validated through neurological investigations. The preliminary results of an EEG-based creativity study are described that use a mix of the sense-making curve analysis paired with EEG data to explore the neurobilogical plausibility of the proposed technique. Next, the most immediate technical directions for interactive machine learning in collaborative drawing are described. Finally, a novel cognitive and software architecture is proposed based on the findings from user studies that may increase the autonomy of the co-creative agent prototype as well as the user's ability to understand and coordinate with such a system.

8.2 Drawing Apprentice System

The Drawing Apprentice system provided a platform to explore what it means for users to collaborate with a co-creative agent in the open-ended artistic domain of improvisational drawing. The work performed in this thesis provides generalizable insights about evaluating co-creative systems in open-ended improvisational domains. In particular, the sense-making curve analysis technique and the user studies we performed on the Drawing Apprentice system provide valuable knowledge about how to effectively evaluate co-creative systems.

Future evaluations of the Drawing Apprentice system (and other simiar co-creative systems) could employ the following features in their experimental design. First, the user should be asked to perform a relatively short (approxiately 3 minutes) non-collaborative drawing task to provide a baseline to quantify the interaction dynamics of the participant's creative process without any intervention. Following this baseline open-ended drawing task, the participant should engage in 3 additional collaboration conditions each lasting approximately 7 minutes:

- **Agent Collaboration:** Collaborate with the co-creative agent.
- **Wizard Collaboration:** Collaborate with a wizard (expert artist) that is controlling the response of the system. Importantly, the wizard should employ their naturalistic collaboration techniques rather than trying to fit their artistic and collaboration style to the system. For example, if artists typically begin their collaborations with 1-line turns and gradually move to synchronous drawing, they should feel free to employ that strategy with the participant, even if the system does not employ that type of procedure.
- **Artist Collaboration:** Collaborate with the artist that was controlling the wizard face to face with pen and paper. Collaborating through an interface drastically reduces the subtle feedback and communication that is mutually available in a face-to-face collaboration. Including this condition provides data about the most naturalistic interaction dynamics of collaborative drawing, resulting in a proper ceiling with which the agent and wizard conditions can be evaluated against.

After each conditions, participants should be asked to engage in a retrospective protocol analysis during which they watch a video of their collaboratin and provide commentary about their experience. During this retrospective protocol analysis, it may be possible for researchers to create a rough sense-making curve with the participant. The researcher could briefly explain the conventions behind the SM curve and ask the participant to provide commentary about their mental state with respect to the SM curve. The SM curve may serve as a tool for mutually understanding and communication about the complex cognitive processes occuring during the creative process.

8.2.1 Learning Object Sequences and Narratives

In addition to performing additional evaluations of the Drawing Apprntice system, there are a number of technical directions that can be explored in the future. Currently, the system recognizes and reasons about individual objects, but its reasoning about the relation (semantic and spatial) is extremely limited. We developed a hard-coded categorical structure within the initial knowledge base of 250 objects. However, as the number of objects grow, a new organizational scheme is required to know which objects should appear together in a meaningful way.

There are multiple ways this information might be learned and encoded. First, we can implement one of the existing semantic/conceptual networks that have a large library of concepts, such as ConceptNet 3. Using this approach, whenever the system learns a new object (through user demonstration and labeling), its label can be used to situated the new knowledge amount other concepts it knows.

Additionally, the system can learn through its experience observing what objects users typically draw together. This narrative module would observe which objects tended to be drawn together, during one individual drawing independently, in collaboration with another human, or in collaboration with the co-creative agent. In the last condition, there should be a feedback mechanism employed before relational knowledge is stored to avoid poisoning the system with bad data coming from the agent. For example, if the agent draws an object that does not semantically or conceptually belong to other objects that have been drawn, it should not be stored as a successful co-occurrence, otherwise the system will learn relational pairings for objects that do not necessarily belong together, according to a human observer.

This type of narrative learning module, would be well suited for learning themes or scenes that occur often in drawings, such as a house scene that might includes a house, tree, person, animal, sun, and clouds. Further, there is spatio-temporal encoded in the construction of these scenes including the order in which the objects are typically drawn as well as the distance and relative positioning of each object relative to each other and the overall composition. This narrative module should be sensitive to all these variables in order to predict what type of object might align with the users current intention, and when and where that object should be drawn.

8.2.2 Partial Object Completion

One of the most requested features from artists was for the system to be able to help them complete objects they began drawing. For example, the user draws a bike tire, and the system helps them draw the frame. There is an algorithm in place in the system designed for this purpose, but there are some unsolved technical issues that need to be addressed before this module can be fully functional.

Once the object blending module has selected the source image from the sketch database to use in the blending procedure, the source and target image (users input lines) are sent to the feature mapping module. This module compares the source and target image to determine which elements from the source image are not present in the target image and therefore marks those as applicable to add to the users target image. This module uses a computational method called histogram of oriented gradients to collect features that are then sent to another existing algorithm called t-SNE on those oriented gradients to align the two images and find common features. Then, the system removes common lines (to avoid redrawing existing features) and finds the remaining elements. These elements are then turned into a series of x and y coordinates that describe lines that can be added onto the canvas.

One significant problem is the resulting lines from the input are not well suited for the drawing module that essentially takes in a series of x,y coordinates and draws a line connecting them. The subtraction process yields many dot-like elements scattered throughout an image. These dots need to be reconstructed into actual lines that can be drawn on the canvas. Further, there needs to be a method for determining which lines (if not all of them) should be drawn during a given turn.

8.2.3 Object Blending

In addition to completing an object in a way that aligns with the users prediction, it could be useful if the agent was able to transform a given group of lines into a new object, i.e. interpret an input as something else. The following procedure could be used to accomplish this within the framework of the current architecture.

Once the object recognition module has classified the users input lines, the label and input lines are sent to the object blending module to determine how to build upon the users input.

When the system is attempting to blend objects, it will select a nearby category to transform the users input into a novel object. For example, at high creativity, the system could decide to turn the users house into a rocket ship. The creative blending algorithms could use nearby classification labels to find sketched images that could be combined with the users input. Then, the selected source image (from the database) is mapped onto the target image (on the canvas) in a similar feature mapping process as described above.

8.2.4 Predictive Drawing

One of the interesting lessons our exploration in co-creative drawing provided was helping to understand how different machine learning approaches yield different constraints and capabilities for a co-creative agent. While the method of data acquisition always affects an agents knowledge, co-creative agents introduce novel training opportunities, such as interactive machine learning, where users train the system in real-time through feedback, such as explicit voting as well as decisions they make in response to the agent (i.e. implicit feedback).

The initial training paradigm was heavily inspired by the enactive theory of cognition in which agents learn through their own experience interacting in situations where they attempt to determine regularities through experimentation in a process referred to as sense-making. In early prototypes, the system could transform user input in a variety of ways. Here, user feedback served to inform the system which type of transformations the user tended to like. However, there are several limitations of this approach.

First, if users do not like any of the transformations the system knows how to perform, it does not matter how much feedback and training it receives, it will not be able to produce responses that make sense to the user. Second, the system learned throughout the course of a drawing, meaning that feedback was gradually ascertained throughout the creative process. Since the agent was gradually tuning its algorithms, the immediate effect of the feedback was diminished, which led users to discount the value of feedback and use it less often.

Third, since the agent emphasizes transforming user input, it did not utilize all of the knowledge potentially available to it through user input. The system could have stored all the user lines and subsequently used those sames lines as output instead of the transformations. This would greatly increase the repertoire of available actions. Then, the system should focus on learning what circumstances each of its different line contributions are most appropriate for given the types of feedback users provide. However, the limitation of learning through interaction would still diminish the immediate effects of feedback, which can mean that a large portion of early interaction with the system is perceived by the user as not effective at collaborating or perhaps even frustrating.

Learning only through direct experience has an additional limitation, namely the assumption that data from the user is a valid dataset to learn from to facilitate collaboration. While expert artists (and expert collaborators) might be effective at providing this data, novices tend to draw similar basic shapes and patterns, which does not necessarily represent the most rich and creative data source.

Through our experimentations, several alternative approaches emerged that we plan to explore in the future. First, we acknowledged the limitations of learning through collaborative interaction, especially with novices, given their limited content knowledge and understanding of what it might mean to collaborate. Second, we acknowledged the need for some base of expert knowledge in order to understand and appropriately respond to users. This lead to the implementation of object recognition, and the drawing algorithms associated with that capability.

Finally, we postulate that it might be more effective to use independent and collaborative human drawing data to train the system in addition to real time interaction. In the proposed machine learning approach, the system would be trained by observing humans drawing, then that knowledge would be utilized during collaboration between the co-creative agent and humans, during which time user feedback would help tune how the available knowledge is actualized for a given user. User feedback would be used to develop a unique user model that constrained how available knowledge was applied to this particular interaction.

This machine learning approach would still be classified as enactive since the agent learns by watching examples and then experimenting with what it learned in those examples through interaction that refines that knowledge through feedback.There are two ways this alternative approach can be implemented: 1) learning from independent expert artists, and 2) learning from collaboration and drawing data from expert artists working together.

8.2.5 Predicting Drawing Behavior of Individual Artist

In this approach, the system would be trained on time-lapse images of one expert artist drawing a complete picture. The granularity of the sampling rate would have to be experimented with, but one could start taking by taking a screenshot of the canvas every 5 seconds. The first step would be to gather enough data to enable the system to develop a basic model for a particular artist. One drawing that lasted thirty minutes would yield 360 time lapse training images with a 5 second sampling rate. Each image would be timestamped and labeled to indicate what drawing it was from and when that particular image occurred in the process. With five such datasets, the system could begin to develop a preliminary model for a particular artists. A similar deep convolutional neural network as the one we used for object recognition could be applied to this scenario to train a model based on the time lapse images collected during the demonstration.

This model would be used to predict the final outcome of an artwork would given some initial user input lines. This model would be queried in real time, as the user draws on the canvas. The feedforward process would happen as follows. First, the user draws a few lines on the canvas. The system takes a screenshot of the canvas that is sent to the neural network to generate a feedforward response that predicts the most likely outcome of the artwork based on what it knows about how other similar drawings have developed in the past. This feedforward output should include both a final product as well as the next step predicted for this particular artwork to take.

A new type of interactive machine learning could be used to fine tune the model.

The system would try to predict, at given time intervals, what the artists next action might be. Then, it would evaluate the actual outcome of the next move and compare it to its prediction and adjust the model based on that feedback. For example, if the artist drew something on a completely different part of the canvas that looked nothing like what the system predicted, this would be considered a large negative feedback event. However, if the user drew something similar to the predicted response in a similar area, this would be treated as a positive feedback event.

None of these processes need to necessarily be revealed to the user as they are engaged in their drawing task. However, once the model can accurately predict what the user might draw next, the system could then generate those lines and apply them to the canvas during a real time collaboration. Generating drawing contributions using this method has the potential to anticipate and assist artists in their creative process. Further, models could be developed for particular artists, which novices could then use to see how experts might contribute to their current drawing.

8.2.6 Predictive Collaborative Drawing

Similar to the predictive drawing scenario, the system could be trained by observing at least two players engaged in a collaborative drawing. This type of neural network would develop an adversarial model, with each player being classified as an adversary. Then, the system would try to predict what a given partner would do in this particular context. Instead of predicting what a particular artist would do next, this model would predict how a given player would respond in a particular scenario.

The training procedure would be similar to the individual drawing scenario outlined above, with expert artists providing some initial training data by completing several collaborative drawings. Here, the training data must distinguish between which players made which lines. Each players line could be a different color when sent as an image to the neural network to make this distinction clear. Additionally, the sampling rate in this case would have to be reduced in order to increase the granularity of the data such that turns can be captured more accurately. A sampling rate of one image per second should suffice in this context.

After the initial model is developed with this preliminary data, the system could then try to predict how each player will respond given the previous players turn as well as the history of interaction (i.e. all previous lines and turns). Then, the players actual response could be used as a feedback mechanism to help refine the model. Finally, a co-creative agent could adopt the model from one of the players for use in a real time collaboration.

8.2.7 Combining Narrative Reasoning with Predictive Drawing

The alternative approaches described here allow for a more full realization of the initial inspiration of the enactive model of creativity. That model characterizes creative cognition as utilizing two distinct modes of cognition, clamped and unclamped cognition. Narrative based reasoning and predictive drawing provide two independent, yet interrelated, reasoning processes that could enable the type of clamping and unclamping observed in human creativity. For example, using the narrative reasoning module, the system could deduce that a user was drawing a city scene and determine objects that are relevant to add and their respective locations. It could use this knowledge to add completely new elements to the drawing in a way that aligns with the conceptual theme of the users artistic intention. Additionally, it could unclamp from narrative reasoning to analyze the individual strokes and predict what type of activity the user might engage in next, such as shading existing structures or thickening lines. From there, the system could clamp onto that task to help the user accomplish activities they are working on. Interaction Dynamics and Coordinating Collaboration

8.2.8 User Experience and Learning From Feedback

There is a tension between creating an autonomous partner that is able to contribute based on its own internal mechanisms and creating a partner that is effective in a collaboration. Part of this problem is the different and often competing motivations of users as they engage with creative technologies. Should the system be viewed as another player with equal autonomy and say as to what stays versus what gets changed?

In order for the user to have enough faith in the system, its knowledge and ability have to meet a certain standard of understanding, predictability, and meaningfulness in its responses. Before this bar is reached, users may quickly become frustrated by a system that is given full autonomy without the proper knowledge to utilize that full autonomy effectively.

The technical approaches identified here can provide additional capabilities for both an inspirational and assistive co-creative AI. Additional control and precision for feedback can greatly improve the assistive use case by allowing artists to define their goals explicitly to the agent. Training the system on expert individual and collaborative drawing can also greatly improve the systems ability to generate responses that truly inspire rather than frustrate the user.

8.2.8.1 Implicit Feedback

There were some general remarks across user groups that related to voting system and way of providing feedback to the system. In general, users wanted the voting and user feedback system to be tied to their implicit drawing behaviors and actions rather than having to provide explicit feedback in the form of binary voting. Many users mentioned manipulating the agents contribution in various ways in a way that could provide feedback to help train the machine learning component of the system. For example, users wanted the ability to scale, move, and remove the agents contributions.

Each of these actions could serve to inform the system about its contribution. For example, if the user drew a tree, and the agent drew a flower, the user might be able to reposition and scale that flower to be more appropriate to the scene they are creating. This feedback would then include multiple pieces of information to help make the narrative module more robust, including information about where flowers typically go relative to trees, as well as what their size ratio is relative to trees.

8.2.8.2 Activity Based vs. Action-Based Feedback

The feedback mechanisms explored in the Drawing Apprentice prototype provide insight into how users might be able to provide training and guidance to a co-creative agent during real time collaboration. In general, users want more immediate control over the agents actions than the current feedback system affords.

The current feedback system is geared toward informing the system how it performed on individual actions, rather than providing feedback about the type of activities in which it has been engaged.This leads to a need to provide frequent feedback with relatively slow response times to that feedback, kind of like steering a large ship. However, if the agents behavior was activity-based rather than action-based, feedback could be used to help direct the types of contributions the agent should make in the near future, rather than what types of actions, overall users tend to prefer.

This type of high-level feedback would be administered less often, leaving the agent essentially on autopilot unless it was performing activities that did not make sense to the user. Explicit feedback should be viewed as a very high impact value since it takes the user out of the creative flow and focuses attention on maintaining the collaboration rather than doing creative activities directly. Implicit feedback should serve to shape individual actions, while high-level direct feedback should be used to guide the overall creative trajectory of the system.

Facilitating this type of activity-based feedback requires a high level communication channel between the agent and user. We found that including a speech bubble describing how the agent interpreted the users actions and informed them of its own intention to be a good initial step in communicating intentions. However, instead of merely informing the user what the agent will do, there should be mechanisms describing the agents plans at multiple levels of abstraction. Further, users should be able to manipulate and modify the proposed plans of the agent. This is especially important to goal-based activities that artists often mentioned as their primary need for collaboration from a co-creative agent.

8.2.8.3 Natural Language Input

The primary goal of the co-creative agent we designed was to facilitate participatory sense-making during collaborative drawing, i.e. co-creating shared meaning structures through interaction. The means of communication available during real time creative improvisation greatly impacts the ability for two agents to coordinate and facilitate participatory sense-making. The speech bubble we created in the prototype moved one step towards a more flexible and robust channel of communication, but there is significant room for improvement.

We observed users engaging in participatory sense-making with the agent by learning to predict what type of contribution the system would make and anticipating its reaction to particular stimuli. Once this understanding was developed, users may eventually engage in a coupled interaction where some activity was established through a participatory sense-making process and is then engaged in. For example, one participant found a particular way of making wavy lines in many different colors. The agent responded with similar wavy lines. After an initial period of experimentation during which the user made sense of the agents reactions. The user was able to engage in a back-and-forth wavy line making activity where the user and system were both engaged in making different (yet similar) types of wavy lines in a variety of colors. Similarly, another user started drawing circles and the system mimicked those circles. Since the system was predictably drawing circles in response, the user and agent were able to engage in a cricle drawing activity.

The machine learning algorithms were designed primarily around action-oriented feedback rather than activity oriented feedback. Thus, when users defined these behavior patterns, they had clear expectations that the system would learn the pattern of what they were doing in the previous turn or last few turns, as opposed to learning overall preferences relevant to the entire drawing session.

Natural language input is one method of interfacing with the system to provide more nuanced information regarding the intention of the user and their expectation for the system. For example, if the user wanted to demonstrate drawing wavy lines or circles, this activity could be verbally labeled, demonstrated, and subsequently engaged in.

Allowing users to clearly delineate what they would like the system to learn or respond to would increase the systems ability to learn relevant information and increase the users ability to predict and understand the systems capabilities. Other verbal comments, such as stating what a group of lines looked like, could also be used to help guide the collaboration. Combining the lasso activity with voice input provides a potentially nuanced means of communication with a co-creative agent. For example, users could circle a region and tell the system they want help shading in the structure that was circled. The lasso tool provides a way for users to point to the system. Deixis is a critically important component of communication and underlies a lot of early language learning. For example, parents often point to an object and then provide the name of that object to children. The child might then point to that object and repeat the label to confirm understanding. Similarly, users might teach the system about objects and activities through pointing using the lasso tool and some natural language accompaniment.

Speech is an intuitive choice in this domain since users could continue drawing while communicating with the system. Humans also naturally use language to communicate during improvisational collaboration. Positive and negative feedback could also be implemented using speech input. Types of speech commands fall under the following categories:

- Labeling an activity
- Labeling an object
- Communicating artistic intention (high-level style or theme and low level tasks)
- Providing feedback
- Giving commands to do an activity or draw a type of object

8.3 Extending Creative Sense-Making Framework

The creative sense-making theoretical framework and coding technique can be applied to more creative domains and types of collaborations to help understand its utility and potentially provide insight into creative collaboration in the target domain. The most immediate application is to apply this coding technique to another dataset we have from a pretend play study between adults and children. Comparing the interaction dynamics between adults and children and between adults can help reveal some differences in the how different leadership styles are represented by the sense-making curves.

Another low hanging fruit to apply the sense-making curve analysis to are interactive museum installation, such as tangible table-top interactions. The proposed technique could quantify the interaction dynamics of multiple participants throughout an extended museum installation by coding video data collected of the exhibit.

We are beginning to explore how the sense-making tool and technique could be used to evaluate the TuneTable, which is an extension of the EarSketch project that tries to teach programming concepts through musical composition and remixing [106,112]. The sense-making curve may provide a rapid and reliable means of evaluating temoprally extended open-ended collaboration on interactive tabletop museum exhibits and interactive installations in general.

8.3.1 Quantifying Creative Sense-Making with EEG

Similar to the gap in research about quantifying the interaction dynamics of open-ended improvisational creativity, there is little cognitive neuroscience work exploring the neural correlates of sense-making in creativity. The work described in this thesis can be extended to compare how the sense-making curve of participants engaged in open-ended creative improvisation compare to their neural activation patterns using electroencephalogram (EEG).

Many studies investigating the neural underpinnings of creativity have focused on finding individual brain regions that are active during particular creative tasks [55] (see [144] for a comprehensive survey of creativity-related neurscience studies and [158] for a review of EEG methods for studying creativity). Researchers have also worked dilligently to document neural pathways and mechaisms recruited for different aspects of the creative process, such as divergent and convergent creative tasks, combinatorial creativity, etc. [43, 67, 157].

EEG studies of creativity focus largely on the internal creative process of individual participants, i.e. charting the neurological basis for the cognitive mechanisms and processes involved in the creative process (in line with the creative cognition approach as described by [156]). However, as described throughout this thesis, the theoretical framework of sense-making introduces a novel aspect to creativity research by describing how meaning, and subsequently creative artifacts, are established through situated interaction and dynamic perception-action feedback loops. This research program points out the importance of including interaction through time as a critical ingredient for creative sense-making.

Instead of focusing on individual creative tasks, or specific ways of being creative (e.g. divergent vs. convergent creativity), the research agenda for creative sense-making focuses on how individuals interact with their environment to dynamically make sense of their creative problem. In these open-ended contexts, there would be many instances of problem finding and problem-solution co-evolution as has been noted in research on design creativity [48, 83, 134]. These types of open-ended interactions would be difficult to quantify using traditional EEG analysis methods.

One large obstacle preventing research into creative sense-making are the current mathematical techniques used for EEG analysis. Neuroscience creativity studies often involve a highly controlled and constrained creative task, which helps researchers filter exceptionally noisy EEG signals in such a way as to enable traditional linear mathematical analysis. Studying open-ended creative activities presents significant challenges for these linear analysis approaches since the 'noisy' elements of the EEG data that is typically filtered out may actually play an important role in the overall neural dynamics recruited during creative sense-making.

We propose a novel chaos-based non-linear EEG analysis technique that allows researchers to study open-ended creative interactions. Instead of investigating the activity of particular brain regions, this approach quantifies the overall coherence of neural signals (i.e. (how much noise is in the signal) as well as the relative source of those signals (i.e. highly localized to particular regions vs. spread broadly across many regions). This analysis approach reveals when quasi-stationary steady states of neural activation occur, which signify a high degree of coherence and localization. This value can be plotted through time and compared to the sense-making curve corresponding to the participant's interaction dynamics throughout their creative sense-making process.

This type of analysis will help reveal what type of sense-making was used by each individual participant and whether individuals, on average, tend to employ sense-making in consistent ways. Sense-making can characterized neurally by delineating between clamped and unclamped cognition. During unclamped cognition (i.e. when one is actively making sense of the premise of a situation) the cognitive agent is considering multiple ways of approaching a problem in their mind, typically without action, until some basic idea is consciously settled upon with some confidence. Once a coherent strategy or plan begins to emerge, cognition is said to be clamped. The directive or plan that was clamped upon tunes ones perception to perceive affordances in the environment that will help facilitate the objective. Early actions help to refine and solidify the plan and clamp cognition further. Once clamped, actions are guided by real time feedback relative to the model (or plan, strategy, directive, etc.) dynamically being developed to achieve a given end.

Transitioning from unclamped cognition to clamped cognition represents the process of sense-making, wherein a cognitive system makes sense of the environment and their current intention to such an extent that it has some idea of what actions to take. Within the creative process, there are many interesting questions to ask with respect to clamping and sense-making, such as:

- How long clamped states last for artists and non-artists?
- Under what conditions do clamped states unclamp, i.e. what were the individuals doing, behaviorally, when their cognition unclamped?
- Do artists have identifiable patterns (either by humans or machine learning) that characterize their unique artistic and cognitive style?

We performed a study applying this new analysis technique on EEG data from an on open-ended drawing tasks. Two tasks were investigated in this study: drawing a visual representation of abstract conepts (e.g. eternity, momentum, inspiration) as well as a completely open-ended drawing session where the participant could draw any image they desired. The sense-making curve technique was employed to qualitative code the behavior and interaction dynamics of the participation as they engaged in the drawing. Preliminary results indicate that when users were coded to be in a 'clamped' cognitive state, their EEG signal entered into a quasi-stationary state, indicating a statistically significant relationship between behavioral markers of clamped cognition and neural activation patterns.

8.3.2 Quantifying Participatory Sense-Making in Collaboration with EEG

The EEG techniques developed can also be applied to analyze how collaboration affects creative sense-making, which introduces a whole host of additional questions relating to social cognition, improvisation, and creative participatory sense-making.

In this experimental setup, there should be a distinction made between novice artists as well as novice collaborators. Presumably those individuals that have experience collaborating will be able to facilitate a more effective collaboration. Thus, pairings for this type of experiment should be sensitive to control for these two variables, introducing conditions that include expert collaborators together, experts with novices, and novices together.

The analysis for this type of study would reveal patterns of sense-making during improvisational interaction and help identify instances of participatory sense-making, i.e. times during which the sense-making of each individual is mutually affecting one another. In particular, interesting questions to ask in this study include the following:

- When and how often do the participants clamp together i.e. EEG signals for both participants corresponded to being in a clamped state, either simultaneously or sequentially?

- How does the graph of each individual person compare to non-collaborative case (either personally, or on aggregate)?

- Is there a pattern or trend in either the individual pattern or relationship between the two patterns for sessions that were subjectively (and objectively) evaluated as a successful collaboration?

8.3.3 Predicting Cognitive Modes to Serve as Creativity Biofeedback

Another interesting application of the proposed EEG approach is using it as a real-time biofeedback mechanisms providing data about the individuals creative sense-making processes. The EEG signals can be fed into a machine learning classifier to detect modes of clamped and unclamped cognition throughout the creative process in real-time, thereby providing a type of biofeedback for individuals to employ during their creative process. As opposed to typical biofeedback signals, such as alpha, beta, gamma EEG values (or some processed combination thereof), this classifier will be uniquely tuned to global shifts in cognition, i.e. changing cognitive modes during the creative process.

There were two data sets collected in the EEG experiments, transformed EEG signals and a resulting unclamped/clamped value, as well as a human generated coding of the observational data into clamped and unclamped states. These two data sets have been shown to be highly statistically related, meaning there is a high probability that they are capturing the same process. The human generated clamping and unclamping code can be used as a ground truth to train a classifier to learn under what circumstances an incoming EEG signal is either in a clamped or unclamped state. Depending on the sophistication of the machine learning approach, the data could either be the transformed fractal dimension value or the raw multi-channel EEG input. Surface learning techniques, such as SVM may be able to classify the fractal dimension signal but not the highly complex and noisy raw EEG signals. Deep learning techniques, such as convolutional neural network may be able to recognizing which visual features of the raw EEG help predict clamped or unclamped, according to the ground truth data.

This machine learning model could be used to generate real time predictions about whether a user (wearing EEG apparatus) is in a state of clamped or unclamped cognition. This prediction could be plotted through time to help them understand and analyze how their cognitive processing has changed throughout their creative process. In particular, this may help individuals identify successful and unsuccessful creative strategies, and become aware when they might need to introduce something new in order to optimally stimulate their creative process.

8.3.4 Using EEG Signals to Control Co-Creative Agents

This real time EEG analysis could also be used to determine the creative strategy of a co-creative agent. For example, if a person has been clamped on a task for an extended period of time, it might benefit their creative process to have something novel introduced. Conversely, if an individual is struggling to maintain a clamped state of cognition, the agent might try to make small contributions to gradually propose new directions the user might be interested in exploring. Then, if the user clamps on a particular idea the agent suggested, that activity could then be pursued by both the agent and the user in tandem.

Collaboration strategies should be derived from data learned in the expert collaboration scenarios. For example, experts might naturally sense when an individual has nearly exhausted their ideas and a general unclamp might occur soon. When an expert senses this, she uses previous experience to detect the signs, including body language and drawing behavior. An agent might be able to achieve similar predictive capabilities by relying exclusively on the EEG signal and its sense-making patterns in combination with drawing behavior.

8.4 Extend Sense-Making Curve Tool

The software tool that was developed to apply the qualitative coding procedure of the creative sense-making framework has a lot of opportunity for improvement. As a first step, the tool can be hosted online and made publicly available. Along with the software, formal documentation describing the coding procedure for producing sense-making curves should be included in the online version. Eight users have extensively used the tool, and based on their informal feedback the following changes can be made to improve the tool.

8.4.1 Keyboard Code Application

Being able to switch code applications using the keyboard would enable coders to more rapidly transition between different states. Some analysts may prefer the slider, but the numerical keyboard input should be introduced as an option to select codes. This feature may reduce the noise generated by the process of changing the slider from one code to another, increasing the overall accuracy of the technique.

8.4.2 Multi-Participant Support

It would be helpful if more than one participants of a video could be coded during a single coding session. For example, in the pretend play study, there were two individuals in the collaboration that needed to be coded. Currently, the tool only supports the creation of a single sense-making curve. Ideally, the tool would support multiple curves per session based on the number of people involved in the collaboration being analyzed. Analysts would still have to re-watch the video each time they created an additional curve, but creating a local repository of curves within the tool would be helpful. This change would involve being able to create multiple sense-making curves for each video and being able to export them with individual file names (or together in a combined excel file). Multi-participant coding would also facilitate data visualization on a per-session basis.

8.4.3 Data Visualization

The tool can implement the matlab mathematical analysis techniques to perform the sense-making curve integrations and other classification techniques within the user interface of the coding tool. The web-visualization environment D3 can be used to visualize the sense-making curve analysis on the web platform. This feature would allow users to code each partiicpant in their data set and immediately visualize the creative trajectory curves and sense-making classifications outlined in Chapter 5. Exploring these charts with different information exploration techniques, such as increasing and decreasing temporal granularity and tuning classification parameters, can help quantify interaction dynamics of open-ended creative collaboration with minimal overhead.

8.4.4 Event Labels

Sense-making curves provide continuous data about interactin dynamics, but they do not contain any information about the type of actions that participants are engaging in through time. This could be solved with a relatively simple technical addition to the tool that allowed users to associate text input (e.g. labels, tags) to different time segments of the sense-making curve. To reduce the amount of time it takes to tag events, the system could employ an autocomplete technique for typing tags as well as present the user with frequently used tags when they begin the event labeling task. This feature could combine the power of event-based coding with the continous coding method of the sense-making curve. With this information, the system could classify interaction dynamic trends as well as quantify the number of particular events and their temporal relationship.

8.4.5 Multi-Video Support

In some experimental setups, multiple cameras are used to capture interaction. In these cases, multiple camera angles may provide more perspectives to help analysts apply qualitative codes to the different participants in an experimental session. Users have requested a dual video stream environment where two angles from the same time periods can be simultaneously presented in the tool (similar to the popular YouTube Doubler tool located at youtubedoubler.com) and coded using one sense-making curve.

8.4.6 Automatic Reliability Assessment

Ensuring that analysts are reliably applying the coding scheme is critical for the success of the proposed technique. Currently, analysts export a file of comma separated values (.csv file) and emply a statistical analysis program called R to perform the inter-rater reliability measure on the .csv files. It may be possible to include this IRR calculation in the tool itself. To fully implement such a feature, the system should support the creation of user accounts that can create projects to which multiple users can contribte. When two users code the same video, there would be an option to compare the reliability of code applications for each participant that has been coded in a session. To maximize the utility of this IRR measure within the tool, the system could perform additional analyses of the curves to identify the time segments were they significantly diverge and present this information to the user in a data visualization. This process can help analysts refine the coding scheme they are using for their target domain and improve overall reliability between analysts.

8.4.7 Automated Coding

It might be possible to train machine learning algorithms to automatically code video data based on training sets comprised of human coded data. This mechanisms could attempt to code the video and inform the user of which segments of the curve it has less confidence in. Then, the user can modify those segments, thereby reducing the overall code application time, while still maintaining accuracy.

The sense-making curve tool allow a rapid and reliable method of quantify interaction dynamics. This tool has many application opportunities in different domains of creative collaboration. Making the improvements listed above may result in a commercially viable tool to help researchers in many fields.

8.5 Conclusions

This chapter described some potential future directions for the creative sense-making research program outlined by this thesis. These future directions follow the three major trajectories explored in this thesis. The creative sense-making theoretical framework and qualitative coding technique can be applied to different open-ended creative collaboration domains to further validate its utility and potentially provide novel insights into the nature of collaboration in these domains. In particular, we described how the technique is being explored for use in evaluating interactive tabletop museum installations. Next, we described our preliminary efforts exploring the neurobiological plausibility of creative sense-making and the sense-making curve. Early results indicate that the basic distinction made between clamped and unclamped cognition in the sense-making curve technique have statistically significant correlations to quantifiable neural patterns. Finally, we described the next steps to improve the user experience of the Drawing Apprentice system.

CHAPTER IX

CONCLUSIONS

The field of computational creativity is beginning to investigate how co-creative agents might interface with the human creative process. These computer colleagues are a mix between creativity support tools helping users achieve creative goals and creative algorithms that generate content autonomously. They have enormous potential because during creative improvisational collaboration, a new form of distributed creativity arises that can lead to emergent, dynamic, and unexpected meaning to support creativity in new ways. However, there is a gap in the literature about cognitive accounts of the interaction dynamics of open-ended creative collaboration, e.g. the rhythm of interaction, style of turn taking, and manner in which participants are mutually making sense of a situation. Such a framework would greatly aid in the design and evaluation of co-creative systems. This thesis began to address that gap by asking the overarching research question: How do humans collaborate in open-ended improvisational creativity, and how can we design co-creative agents to achieve similar benefits as human collaboration?

The thesis statement was that co-creative agents, such as collaborative drawing partners, can inspire new ideas and motivate interaction during open-ended and improvisational creative collaboration on a shared canvas. Furthermore, I claimed using participatory sense-making as a theoretical lens to model and quantify co-creation (e.g. interaction dynamics, emergent meaning, coupling, autonomy) can help objectively evaluate the effectiveness of interaction designs and technical approaches in co-creative systems. A mixed-method approach was employed to explore this claim, including: empirically investigating open-ended improvisational collaboration, developing techniques for computationally modeling interaction dynamics, and building a co-creative drawing partner as a technical probe to explore how the proposed framework can benefit the design and evaluation of co-creative systems.

My personal motivation for investigating this domain was my extensive practice-based experience in the domain of collaborative drawing over the past 15 years. During my experiences collaborating with many individuals of different experience levels, I consistently saw how turn-taking based open-ended collaboration can benefit both experts and novices, such as:

Lower barrier of entry: Novices seem to feel more comfortable engaging in the artistic creative process when they are only expected to draw a few lines per turn in a collaboration. The subjective value attached to any individual line in a turn is reduced due to the highly mutable state of the artwork over the course of many turns. As a result of this value change, the experience shifts from trying to represent a particular idea (which is difficult) towards actively participating in a gradual unfoldment of evolving ideas (which is easier).

Creative inspiration: Each contribution inspires and invites the participant to respond to it, making an addition or possibly resolving some tensions that arose when an expectation was violated (thus inspiring new ideas). Often, the end result is more creative than what could have been accomplished individually due to the diversity of ideas, skills, styles, and artistic visions.

Motivation to continue: Collaboration motivates individuals to continue by shifting the value of engaging in art-making from: A) producing a good final product, to B) making interesting contributions in response to your partner through time. The dialogical element of the collaboration seems to compel the interaction to move forward, like participating in a conversation that naturally unfolds through time.

Given these experiences and an understanding of the academic landscape of computational creativity, I set out to design, develop, and evaluate a co-creative drawing partner. However, evaluating the success of the system became problematic. It was difficult to precisely formulate the ultimate goal of such a system in the context of creativity and cognitive science research since there was not a unifying cognitive theory or method for studying and evaluating creative collaboration in general. As a result of this gap in the research, we began using the theoretical principles of enaction and participatory sense-making (e.g. dynamic and emergent meaning, and dynamic feedback) to inform the design of the Drawing Apprentice system. However, it was difficult to tease apart how each individual design decision and system limitation were influencing the creative engagement and quality of the collaboration given the open-ended nature of collaboration and number of confounding variables that could be influencing decisions and behavior.

I began the effort of mapping the ideas of enaction and participatory sense-making to a computational framework, qualitative coding technique, and analysis approach that could inform the design and evaluation of co-creative systems. Once I realized that this framework had the potential to be much more impactful than building an individual instance of a co-creative system, the focus of the thesis shifted more toward building the theoretical framework of creative sense-making and working to evaluate its productivity through applying its ideas and analysis techniques to multiple domains of open-ended creative collaboration and publishing the results to gauge whether the community found this perspective valuable. Through these efforts in theory building, a technique was developed to code qualitative data of open-ended interactions in such a way as to allow mathematical analysis of interaction dynamics using continuous functions (e.g. integrations, moving averages, etc.) to quantify how participants were engaged in sense-making and participatory sense-making. The specific research questions asked in this thesis were:

RQ1: *Can a co-creative drawing partner leverage user input and feedback to facilitate participatory sense-making in a similar manner as humans, e.g. emergent meaning, coupled interaction, inspiring dialogical interaction?*

The answer to this question is partially, but that is due partly to humans being particularly well skilled at attributing intentionality. There are severe limitations to relying solely on current user input and session based training data as knowledge sources for a co-creative agent as was initially proposed for the Drawing Apprentice. Skilled collaborative artists can certainly use that system to inspire their creativity if they learn to understand the constraints of the system, and it can offer value in that respect, but that is not generalizable to novices necessarily. Later architectures for Drawing Apprentice included a knowledge based of 20k human sketched objects, and a sketch recognition module that could classify the users sketch in real time, which increased user engagement and created an object based drawing dialogue between the user and system. Next, the system requires a collaboration module that is trained on human collaboration data to teach it when, what, and why it should draw an object or feature of an object during a collaboration and why.

Machine learning should be applied at different levels in a co-creative agent, and much attention needs to be paid to training data for different types of skills the agents require to facilitate co-creation. For example, the agent needs different knowledge for knowing what the user is drawing, what it should draw, and why it should draw that particular thing at that point in time. Without some attempt at each of these reasoning processes, users can become bored and frustrated with the system.

With this question, we worked to understand what types of information and interactions users employ to understand co-creative agents and make sense of this novel collaboration context. Here, the systems ability to draw is less important than the interface elements meant to contextualize and explain the agents contribution. The user study focused on the interface design helps address this question by evaluating what types of mental models users construct of the system and how the interface elements and interaction design effect that conceptualization. Feedback from public demonstrations also provides a unique perspective on this question because users have to rapidly orient themselves to the system during their brief interaction with the system.

At its core, this question seeks to understand to what extent our current implementation of the Drawing Apprentice is able to achieve meaningful collaboration, and in particular meaningful as being defined as participatory sense-making. While there are obvious limits on how well a co-creative agent can seamlessly integrate into a creative collaboration, here we seek to understand whether our technical approach lends itself to facilitating participatory sense-making or the shared and emergent construction of meaning through multiple coordinated interactions. Participatory sense-making is critically important to collaboration because it helps individuals move beyond their creative boundaries by introducing new ideas, essentially extending the zone of proximal development by challenging users to engage with new artistic ideas. However, there is a significant difference between pushing creative boundaries and frustrating the user, and this changes for different user groups. Our findings indicate that meaningful collaboration changes depending on what user group is exposed to the tool. In particular, artists and non-artists have much different conceptions of what a good collaborative partner would be. The public demonstrations and art gallery exhibitions directly answer this question by providing insights into these user groups and how they would ideally like co-creative agents to interface with their creative process.

RQ2: *Can the cognitive science theory of participatory sense-making be productively mapped to the field of co-creation to provide a solid cognitive paradigm and framework to understand the unique sociocognitive elements of co-creation?*

The quality of thick descriptions and conceptual understanding of collaboration published in the field of computational creativity (e.g. ACM Creativity and Cognition and International Conference on Computational Creativity) using this approach indicate the answer to this question is yes, but the theories of enaction and participatory sense-making were underspecified with respect to creativity and collaboration in particular. The extensions described in the text offer a way to bridge this gap and demonstrate the value of doing so, i.e. more informed and targeted findings about design decisions of co-creative systems. In particular, the clamping theory and enactive model of creativity offer two missing pieces to help understand how cognitive states fluctuate dynamically through time during creative sense-making. While these models still require further empirical investigation and validation, their conceptual structure has enabled further nuanced discussion in the field of computational creativity about interaction dynamics regarding co-creation.

RQ3:*Does continuous data quantifying the interaction dynamics of co-creation help evaluate the technical approaches and interaction design of co-creative systems?*

The type of findings and data presented in this thesis indicate the proposed creative sense-making framework can be productively used to ask fine grained questions about the user experience, creative process, and style of collaboration–both between humans and co-creative systems and between humans in general. Having a continuous curve describing the functional role of an agents actions in a continuous sense-making process enables new types of analyses that capture the coordination and interrelatedness of the collaborators through time.

The new creative sense-making framework and coding technique was applied to both human collaboration data and user study data from the co-creative system. The analysis showed distinct differences between human collaboration and human-computer collaboration with the Drawing Apprentice system, and the findings provided ideas about how to make the collaborative experience more human-like. While the Drawing Apprentice system might not collaborate with users quite like a human yet, the proposed framework and evaluation technique can now clearly quantify its shortcomings and inform future design directions for the Drawing Apprentice system and other co-creative systems.

In summary, this thesis extended the cognitive science theory of enaction and participatory sense-making to the domain of open-ended creative collaboration to formalize this theory in computational models of creative collaboration. This knowledge primarily contributes to the fields of computational creativity, human-computer interaction, cognitive science, and creativity research. A new conceptual framework and accompanying analysis technique called creative sense-making was proposed as a new method to visualize and quantify the interaction dynamics of creative collaboration, e.g. the rhythm of interaction, style of turn taking, and manner in which participants are mutually making sense of a situation. This is critical from an enactive perspective since this theory proposes that meaning emerges in two primary ways: through the dynamics of the interaction as well as the content of individuals actions within the overall flow of actions. The creative sense-making data analysis technique and tool provide a new method to reliably and rapidly quantify interaction dynamics such that they can be mathematically analyzed using continuous functions (i.e. moving averages, integrations) to understand how collaborations flow through time (versus the typical discrete, event-based qualitative coding). This technique may be generalizable to other fields studying and designing products for open-ended human interactions. Finally, this thesis produced a web-based co-creative drawing agent that learns through interaction and can serve as an experimental platform for studying different techniques of interactive machine learning, human-computer collaboration, and human-human collaboration.

APPENDIX A

PRETEND PLAY CODING SCHEME

This appendix describes the detailed behavioral markers used to code the pretend play video data according to creative sense-making technique using sense-making curves. This coding scheme describes the behavioral markers for determining whether participants are clamped or unclamped, including the degree and direction of the unclamp. The numerical codes used by the analysts are based on a 0-4 scale. These numbers each have a direct mapping to the 1 to -1 scale used in the sense-making curves themselves. This numerical transduction was done at the request of analysts to help mitigate the switching time between codes that often occur in rapid succession. Additionally, this transduction separate the curve the analysts saw from those used during the analysis, providing a layer of obfuscation between the analysts and the predictions of the creative sense-making theory.

- Behaviors coded as a 0 value (later transduced to -1)
  - Mostly disengaged
  - Not trying at all
  - Arms crossed/ in pockets
  - Hands by side
  - Looking at the experimenter confused
  - Blank stare
- Behaviors coded as a 1 value (later transduced to .5)
  - Searching through the toy box

- Moving around the play space
- Looking at objects before placement and in between arrangements
- Gathering objects from different places on mat (similar to searching in the box)
- Re-orienting ones position on the mat, e.g. moving to a different location

• Behaviors coded as a 2 value (later transduced to a 1)

- Place items, arranging items, building something up
- Stop playing to grab another object
- Restructuring the space outside the story/action performances
- Picking up another character for the play activity
- Restructure the character and/or space, i.e. building a composite object in between actions, re-arranging the how an object is constructed
- Re-arranging characters, possibly for future use
- Testing out uses of toys
- Meta-communication and instructions related to gameplay
- Clearing the matt off
- Combining objects to make a new character

• Behaviors coded as a 3 value (later transduced to -.5)

- Holding character
- Waiting to act while still embodying character
- Active stance towards the mat
- Leaning forward
- Brief pauses and hesitations

  - If participant is playing, then takes a break to laugh, that break is a 3

- Behaviors coded as a 4 value (later transduced to 0)
  - Performing embodied action with character (all levels of intensity)
  - Talking in character, i.e. diegetic communication
  - Embodied play actions during setup is still a 4
  - Fluid play actions in general (except for hesitations between actions)

APPENDIX B

DRAWING APPRENTICE USER STUDY DETAILS

B.1 Summary

This appendix describes the details and procedure of the Drawing Apprentice creativity study, including the protocol and script, the survey tool and results, and the adaptation of the sense-making curve behavioral coding scheme from pretend play. This domain translation of the sens-making curve coding scheme shows the power of the creative sense-making coding and analysis technique for domain independent analysis of interaction dynamics in open-ended creative collaboations.

B.2 Drawing Apprentice Creativity Study Protocol

The following is a script for researchers conducting the creativity study for the Drawing Apprentice system.

B.2.1 Greeting Participant

Thanks for agreeing to participate in our study, we really appreciate your time. Were researching artistic creativity and we would like to observe how you interact with a drawing system. In the study, you will be asked to create two drawings and describe your experience afterwards. Each drawing task will last about 20-25 minutes, and the whole experiment should take about an hour. Its important to remember there are no right or wrong answers, and you dont have to have any artistic experience to participate in the study. You can stop the experiment at any point for any reason.

If you agree to participate, well ask you to fill out a brief demographic survey about your previous art experience, and then you can begin the drawing tasks. Each drawing lasts 12 minutes, and afterwards well watch a video of your drawing together.

Ill ask you to describe what you were thinking about during the drawing as we watch the video. Then, you will be asked to fill out a survey about your experience and answer a few questions. All your data will be anonymized.

[Obtain consent]

[Demographic survey] Before we begin, could you please fill out this short survey?

[Open the DApp System]

B.2.2 Task Instructions

The systems youll use today are collaborative drawing programs, meaning the systems will collaborate and draw with you. The interfaces for the two drawing tasks are the same, but the algorithms of each system work differently. Both systems use artificial intelligence to learn how to draw by observing and reacting to your lines and feedback. Ill show you how it works now.

When you stop drawing for more than about 2 seconds, the system will begin drawing. You can change the color of your lines and the systems lines with these color selectors on the top of the screen. When you change the line color, that changes both the systems line color as well as your line color. You can help teach it what types of lines you like by using the voting buttons on the right side of the screen after the agent finishes its turn. The systems creativity settings, which affect how it reacts to you and the types of lines it creates, can be adjusted using the creativity slider on the left side of the screen.

Please spend the next 12 minutes collaborating with the Drawing Apprentice system. The system is designed for abstract drawing, but you are free to draw whatever youd like. Ill let you know when time is up.

B.2.3 Retrospective Protocol Analysis

Turn on voice recorder

Stop camtasia recorder

Either: watch camtasia, or video camera

Ask: What were you thinking here? Can you expand a little on your thoughts at this point?

B.2.4 Survey Tool

The IRIS UX Survey instrument (located at: http://www.ux-tool.com/index.php) was utilized to evaluate creative engagement and the overall user experience of the tool in the two conditions.

B.2.5 Semi-Structured Interview

Did your drawing behavior affect the systems drawing behavior? How do you know? How did the system affect your artistic creative process? How did your interaction with the system change over time? Did you use the voting buttons and creativity slider? If so, what were their effect?

Debrief What was similar and different about your interaction with the system in the two drawing tasks?

B.3 Drawing Apprentice Creativity Study Survey Tool

This section shows the IRIS UX survey employed in the Drawing Apprentice creativity study. In the pilot study, we explored the use of a custom survey deployed on Google Forms as well as the Creativity Support Index [17]. The creativity support index had limited questions about collaboration and aesthetics. We decided to the use a customized version of the IRIS UX sruvey as it focused on the experience of the tool as well as usability issues. However, we found that the survey tool overall was not as informative as we had hoped. Part of this reason is that the particular questions we asked could have been more focused on emergent and dynamic meaning, dialogical interaction, and collaboration styles that were discovered after analyzing the qualitative data. Another explanation is that the survey measured creative engagement with the novel activity of collaborative drawing. However, since the difference between the two conditions is not as great as the difference between non-collaborative drawing and collaborative drawing, it is difficult to tease apart the effectiveness of the agent using this data collection tool. In the UI study, a survey tool was used as a boundary object to facilitate discussion with the participants. For example, they chose between the two conditions on various metrics, then explained why they choose one version over another. This seemed to be a more effective method for investigating the users experience of various conditions. The survey questions and results are depicted below. Task 1 is the wizard of oz collaboration, and task 2 is the agent collaboration (in the results).

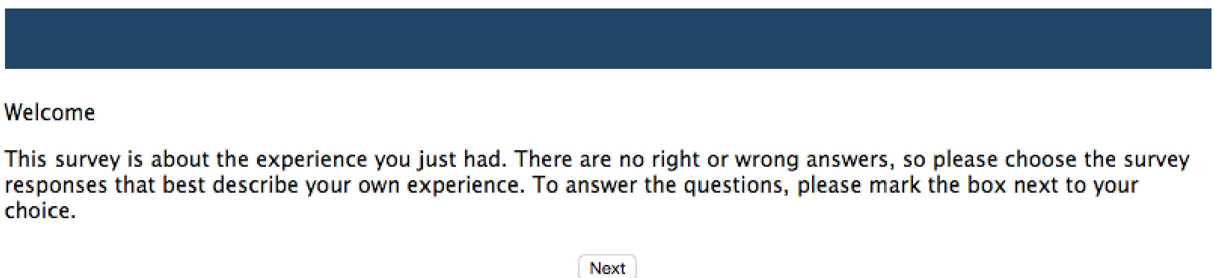


Figure 49: Introduction screen to the survey tool.

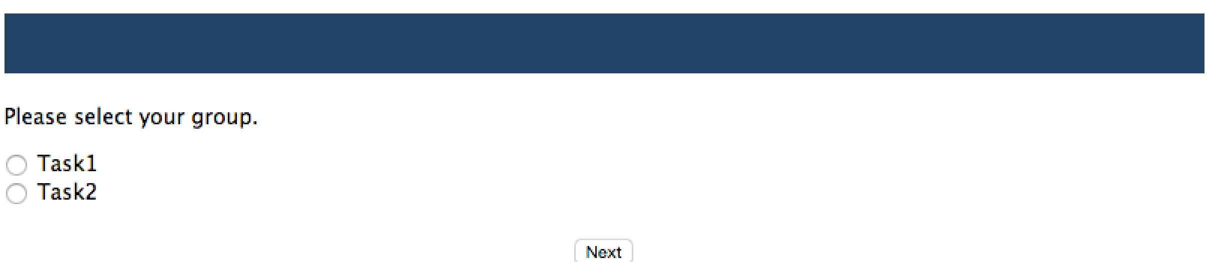
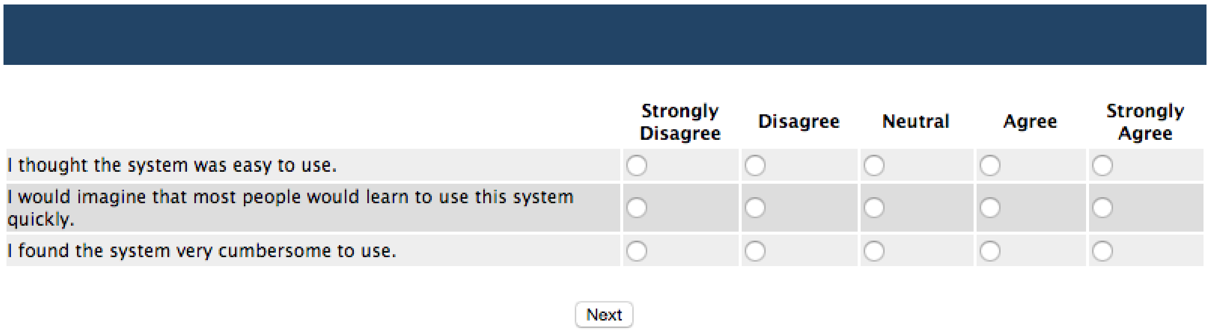


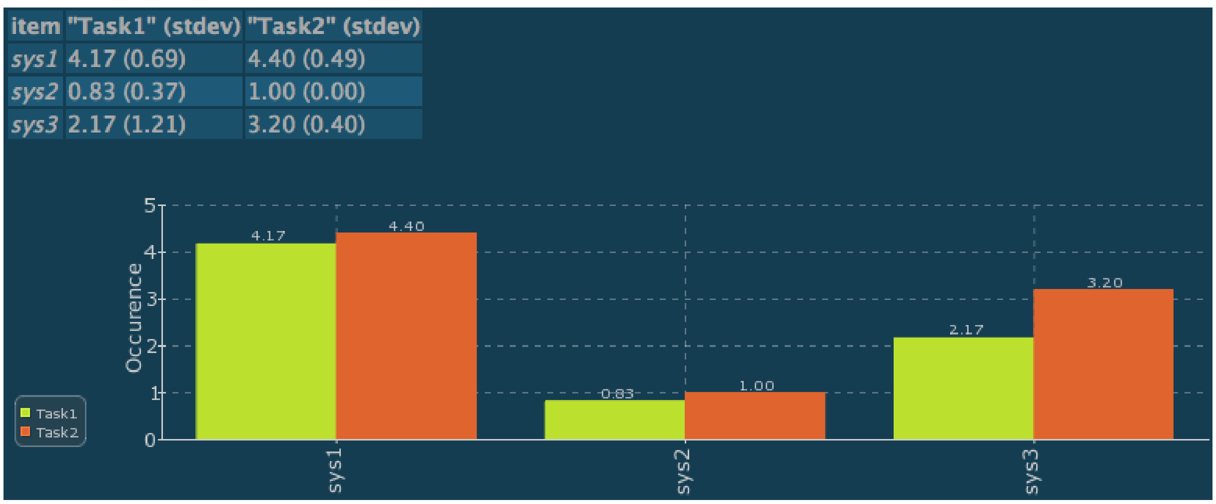
Figure 50: Survey question to differentiate between conditions, e.g. agent collaboration or wizard of oz collaboration.



	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I thought the system was easy to use.					
I would imagine that most people would learn to use this system quickly.					
I found the system very cumbersome to use.					

Next

Figure 51: Usability questions on the survey.



item	"Task1" (stdev)	"Task2" (stdev)
sys1	4.17 (0.69)	4.40 (0.49)
sys2	0.83 (0.37)	1.00 (0.00)
sys3	2.17 (1.21)	3.20 (0.40)

Figure 52: Usability results from the survey.

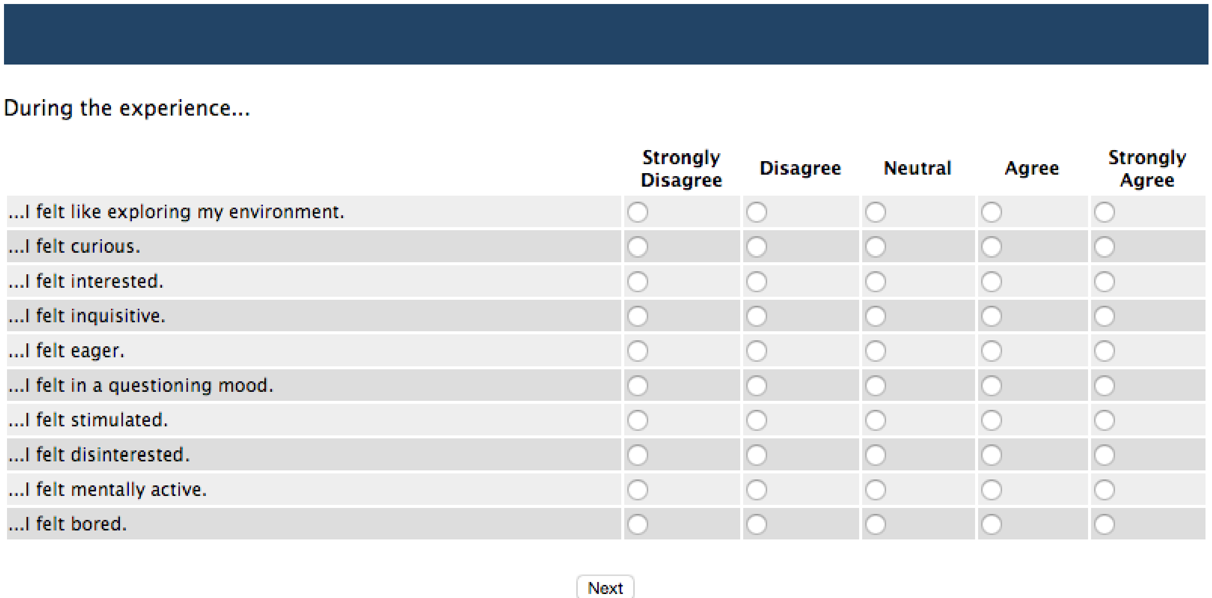
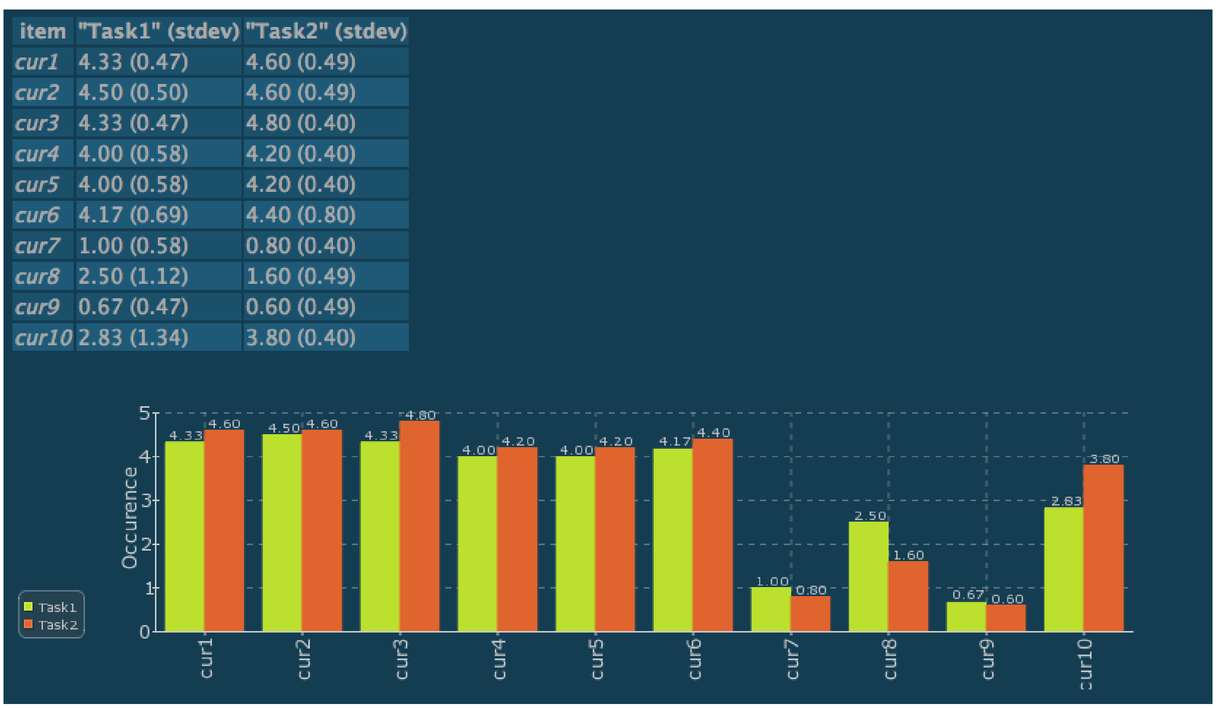
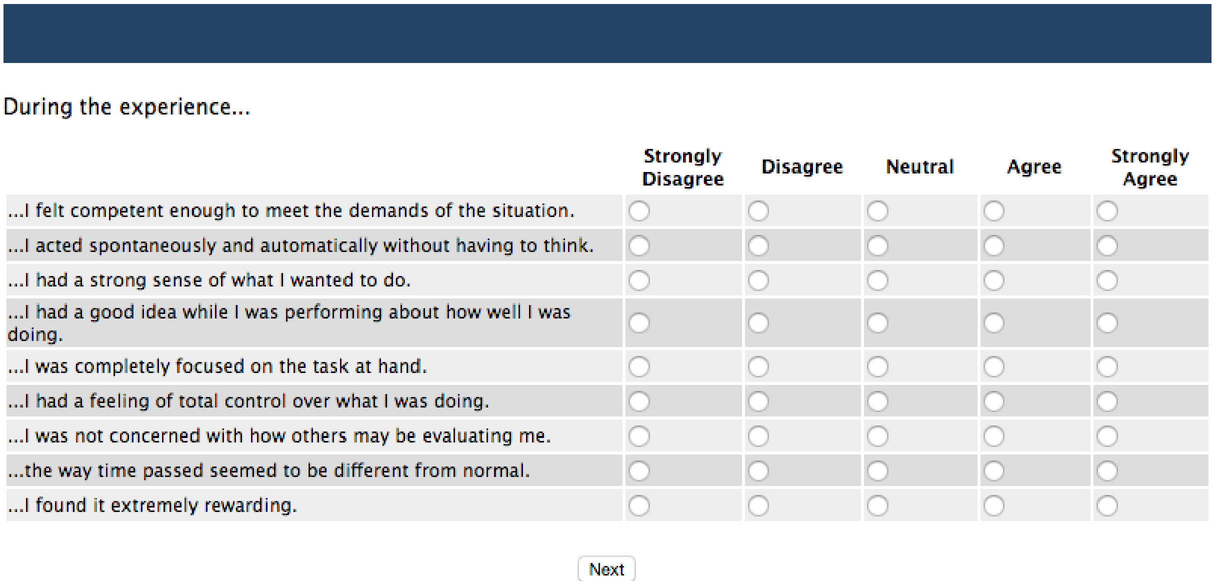


Figure 53: Survey questions related to overall engagement with the tool.



item	"Task1" (stdev)	"Task2" (stdev)
cur1	4.33 (0.47)	4.60 (0.49)
cur2	4.50 (0.50)	4.60 (0.49)
cur3	4.33 (0.47)	4.80 (0.40)
cur4	4.00 (0.58)	4.20 (0.40)
cur5	4.00 (0.58)	4.20 (0.40)
cur6	4.17 (0.69)	4.40 (0.80)
cur7	1.00 (0.58)	0.80 (0.40)
cur8	2.50 (1.12)	1.60 (0.49)
cur9	0.67 (0.47)	0.60 (0.49)
cur10	2.83 (1.34)	3.80 (0.40)

Figure 54: Survey results from overall engagement with the tool.

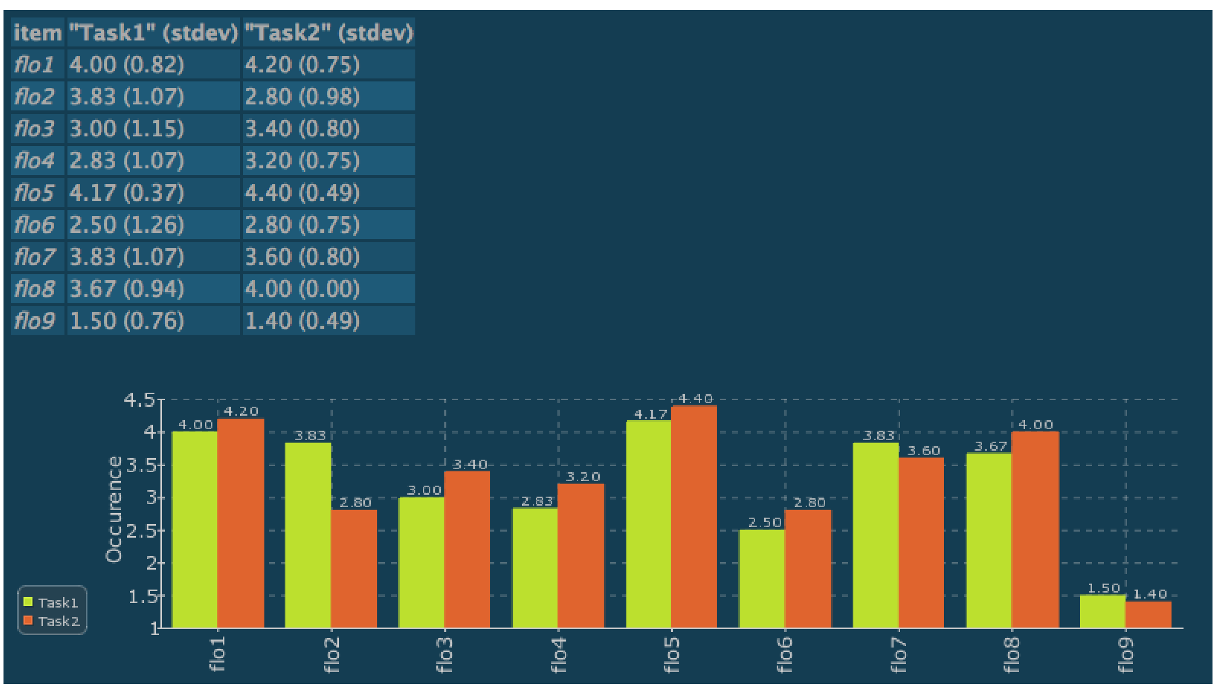


During the experience...

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
...I felt competent enough to meet the demands of the situation.					
...I acted spontaneously and automatically without having to think.					
...I had a strong sense of what I wanted to do.					
...I had a good idea while I was performing about how well I was doing.					
...I was completely focused on the task at hand.					
...I had a feeling of total control over what I was doing.					
...I was not concerned with how others may be evaluating me.					
...the way time passed seemed to be different from normal.					
...I found it extremely rewarding.					

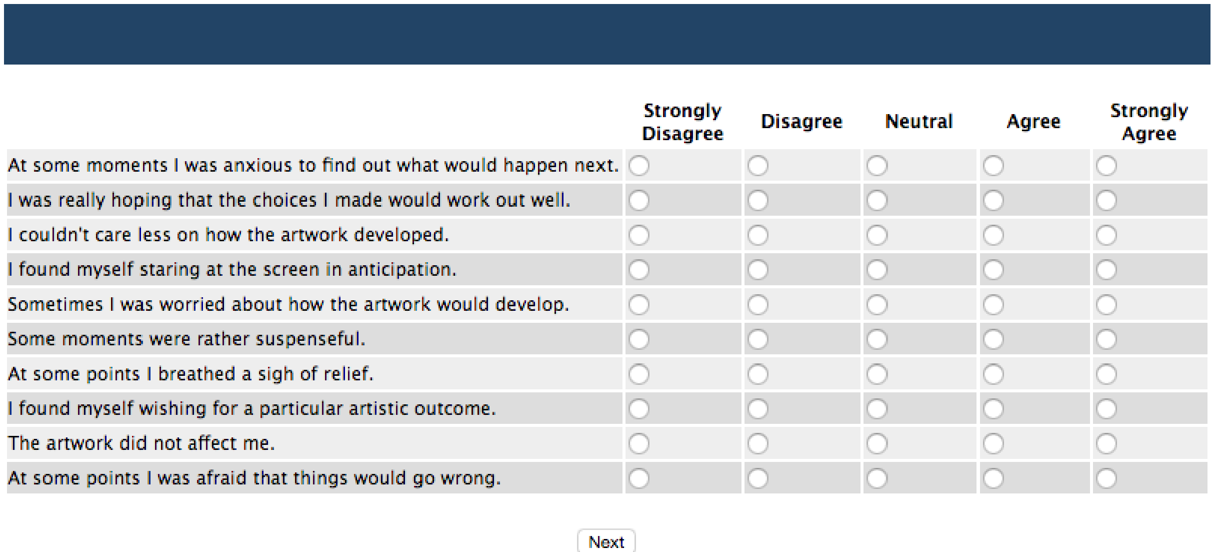
Next

Figure 55: Survey questions related to creative flow.



item	"Task1" (stdev)	"Task2" (stdev)
flo1	4.00 (0.82)	4.20 (0.75)
flo2	3.83 (1.07)	2.80 (0.98)
flo3	3.00 (1.15)	3.40 (0.80)
flo4	2.83 (1.07)	3.20 (0.75)
flo5	4.17 (0.37)	4.40 (0.49)
flo6	2.50 (1.26)	2.80 (0.75)
flo7	3.83 (1.07)	3.60 (0.80)
flo8	3.67 (0.94)	4.00 (0.00)
flo9	1.50 (0.76)	1.40 (0.49)

Figure 56: Survey results related to creative flow.



	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
At some moments I was anxious to find out what would happen next.					
I was really hoping that the choices I made would work out well.					
I couldn't care less on how the artwork developed.					
I found myself staring at the screen in anticipation.					
Sometimes I was worried about how the artwork would develop.					
Some moments were rather suspenseful.					
At some points I breathed a sigh of relief.					
I found myself wishing for a particular artistic outcome.					
The artwork did not affect me.					
At some points I was afraid that things would go wrong.					

Next

Figure 57: Second round of questions about sustained creative engagement in the survey.

item	"Task1" (stdev)	"Task2" (stdev)
sus1	4.50 (0.50)	3.60 (0.80)
sus2	0.83 (1.07)	1.40 (0.49)
sus3	2.50 (1.12)	1.80 (0.40)
sus4	4.33 (0.47)	3.80 (0.98)
sus5	3.83 (1.07)	3.20 (0.98)
sus6	4.00 (0.58)	3.40 (0.80)
sus7	2.67 (1.37)	2.40 (1.02)
sus8	1.17 (0.90)	1.00 (1.10)
sus9	1.83 (1.07)	2.60 (0.49)
sus10	1.17 (1.07)	2.80 (0.40)

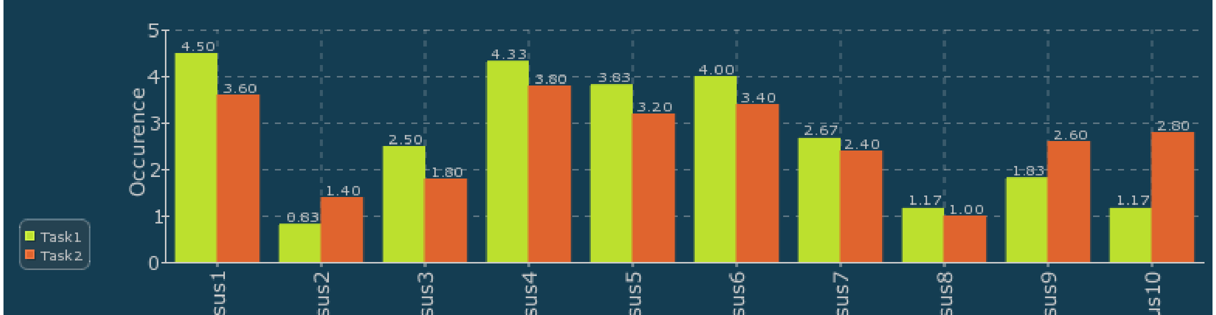
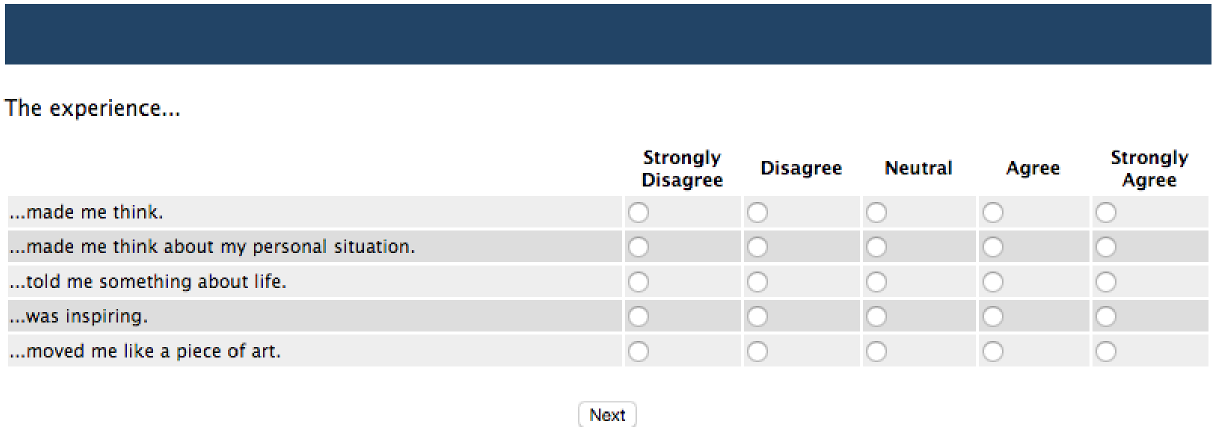


Figure 58: Survey results related to sustained creative engagement.

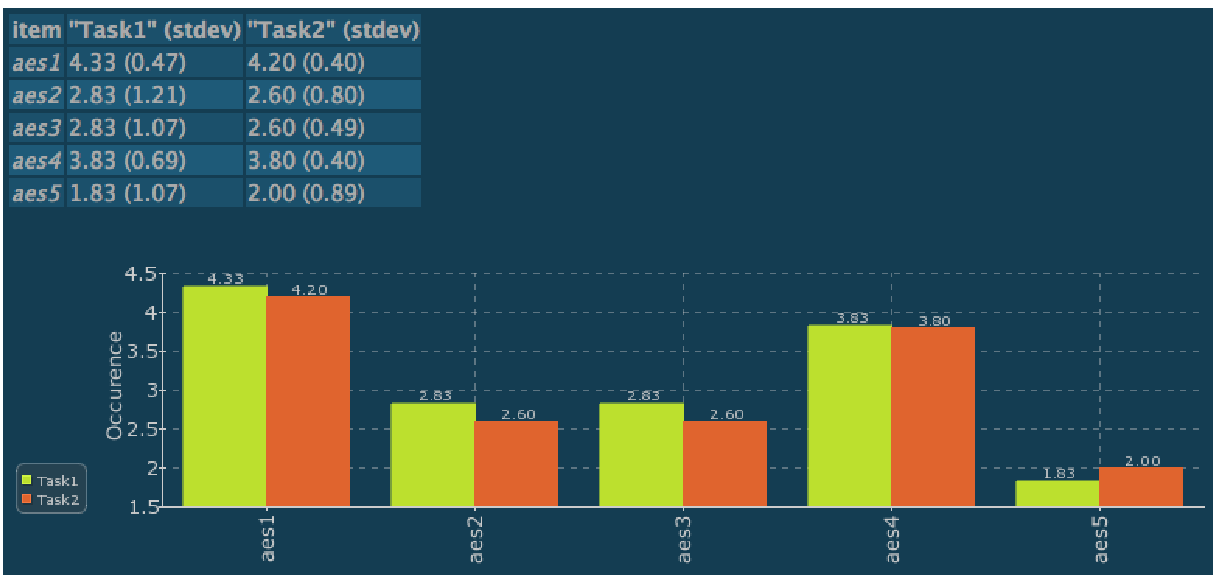


The experience...

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
...made me think.					
...made me think about my personal situation.					
...told me something about life.					
...was inspiring.					
...moved me like a piece of art.					

Next

Figure 59: Questions related to how meaningful the collaboration experience was in the survey.



item	"Task1" (stdev)	"Task2" (stdev)
aes1	4.33 (0.47)	4.20 (0.40)
aes2	2.83 (1.21)	2.60 (0.80)
aes3	2.83 (1.07)	2.60 (0.49)
aes4	3.83 (0.69)	3.80 (0.40)
aes5	1.83 (1.07)	2.00 (0.89)

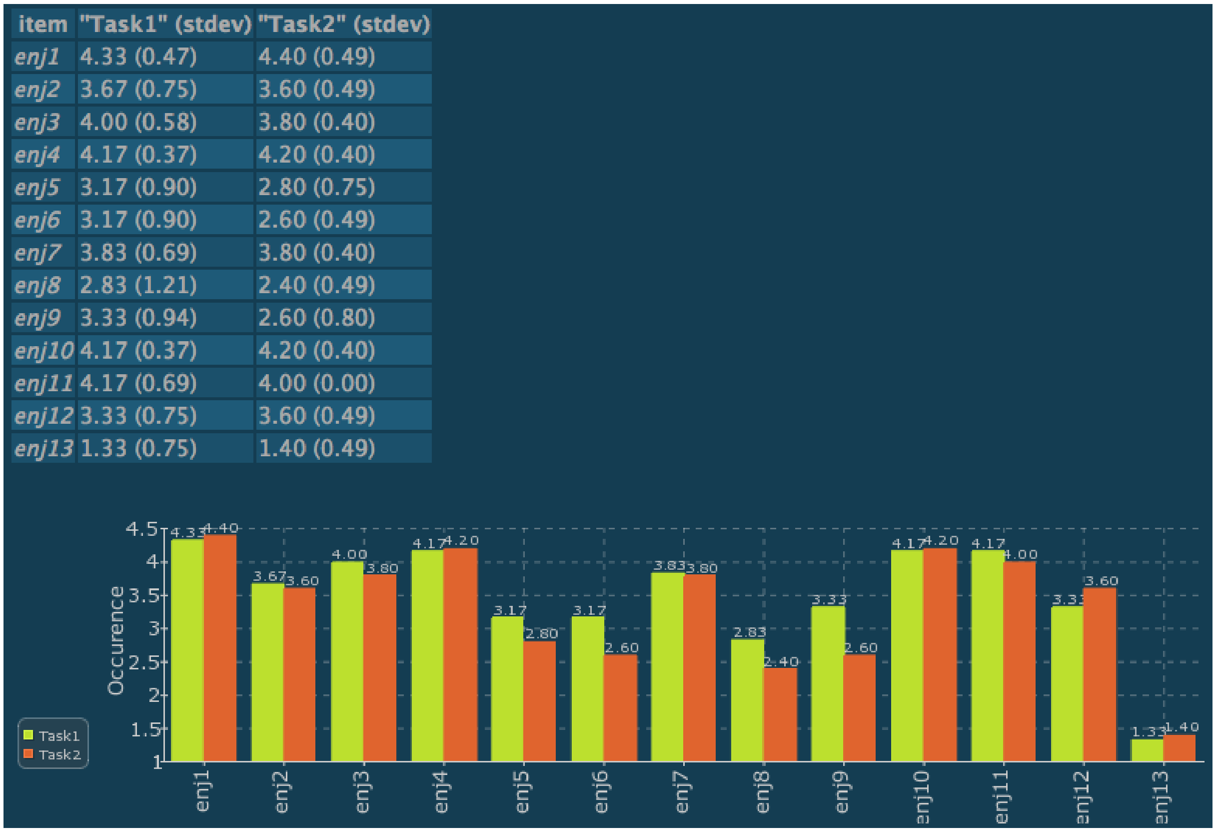
Figure 60: Survey results about how meaningful the collaboration experience was.

The experience...

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
...was pleasant.	○	○	○	○	○
...was gratifying.	○	○	○	○	○
...was rewarding.	○	○	○	○	○
...was amusing.	○	○	○	○	○
...was exhilarating.	○	○	○	○	○
...was thrilling.	○	○	○	○	○
...was exciting.	○	○	○	○	○
...was melancholical.	○	○	○	○	○
...was moving.	○	○	○	○	○
...was appealing.	○	○	○	○	○
...was pleasing to the senses.	○	○	○	○	○
...made me feel proud.	○	○	○	○	○
...made me feel competent.	○	○	○	○	○

Next

Figure 61: Questions related to the affective response and overal evaluation of the experience on the survey.



item	"Task1" (stdev)	"Task2" (stdev)
enj1	4.33 (0.47)	4.40 (0.49)
enj2	3.67 (0.75)	3.60 (0.49)
enj3	4.00 (0.58)	3.80 (0.40)
enj4	4.17 (0.37)	4.20 (0.40)
enj5	3.17 (0.90)	2.80 (0.75)
enj6	3.17 (0.90)	2.60 (0.49)
enj7	3.83 (0.69)	3.80 (0.40)
enj8	2.83 (1.21)	2.40 (0.49)
enj9	3.33 (0.94)	2.60 (0.80)
enj10	4.17 (0.37)	4.20 (0.40)
enj11	4.17 (0.69)	4.00 (0.00)
enj12	3.33 (0.75)	3.60 (0.49)
enj13	1.33 (0.75)	1.40 (0.49)

Figure 62: Survey results about the affective response and overal evaluation of the experience.

B.4 Drawing Apprentice Creative Sense-Making Coding Scheme

Below is the creative sense-making coding scheme for the users as they engage with the Drawing Apprentice application. Coding the agent's behavior is much more simple since there is only a binary drawing/not drawing distinction, meaning it is coded as either a (4- drawing) or (0- inactive).

- Behaviors coded as a 0 value (later transduced to -1)
  - Completely disengaged
  - Confused looking
  - Leaning back from the display

  - Lowered pencil/stylus
  - Gaze away from screen
  - Distracted

- Behaviors coded as a 1 value (later transduced to .5)
  - Providing explicit feedback through voting
  - Verbalizations and commentary, e.g. laughing, chuckling, sighing
  - Moving and resituating the device

- Behaviors coded as a 2 value (later transduced to a 1)
  - Changing the color, thickness, or settings
  - Physically moving around to get a different view
  - Performing simulation drawing actions above device

- Behaviors coded as a 3 value (later transduced to -.5)
  - Pausing after completing a turn and waiting attentively to see what happens
  - Pauses in-between drawing actions

- Behaviors coded as a 4 value (later transduced to 0)
  - Engaged in drawing action currently

REFERENCES

[1] Augello, A., Infantino, I., Pilato, G., Rizzo, R., and Vella, F., "Combining representational domains for computational creativity," in *Proceedings of 5th international conference on computational creativity, ICCC 2014, June*, 2014.

[2] Auvray, M., Lenay, C., and Stewart, J., "Perceptual interactions in a minimalist virtual environment," *New ideas in psychology*, vol. 27, no. 1, pp. 32–47, 2009.

[3] Aydede, M. and Robbins, P., *The Cambridge handbook of situated cognition.* Cambridge University Press, 2009.

[4] Bae, S.-H., Balakrishnan, R., and Singh, K., "Everybodylovessketch: 3d sketching for a broader audience," in *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pp. 59–68, ACM, 2009.

[5] Batey, M. and Furnham, A., "Creativity, intelligence, and personality: A critical review of the scattered literature," *Genetic, Social, and General Psychology Monographs*, vol. 132, no. 4, pp. 355–429, 2006.

[6] Baylor, A. L., "Promoting motivation with virtual agents and avatars: role of visual presence and appearance," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 364, no. 1535, pp. 3559–3565, 2009.

[7] Beňuš, Š., "Variability and stability in collaborative dialogues: turn-taking and filled pauses," in *Tenth Annual Conference of the International Speech Communication Association*, 2009.

[8] Bilda, Z., Gero, J. S., and Purcell, T., "To sketch or not to sketch? that is the question," *Design studies*, vol. 27, no. 5, pp. 587–613, 2006.

[9] Biles, J. A., "Interactive genjam: Integrating real-time performance with a genetic algorithm," in *Int. Computer Music Conf.(ICMC98)*, pp. 232–235, 1998.

[10] Biles, J. A., "Genjam: Evolution of a jazz improviser," *Creative evolutionary systems*, vol. 168, p. 2, 2002.

[11] Boden, M. A., *The creative mind: Myths and mechanisms.* Routledge, 2003.

[12] Bown, O., "Empirically grounding the evaluation of creative systems: incorporating interaction design," in *Proceedings of the Fifth International Conference on Computational Creativity*, pp. 112–119, 2014.

[13] BROOKS, R. A., "Intelligence without representation," *Artificial intelligence*, vol. 47, no. 1, pp. 139–159, 1991.

[14] CAILLOIS, R., "Man, play and games: University of illinois press," 2001.

[15] CANDY, L., "Computers and creativity support: knowledge, visualisation and collaboration," *Knowledge-Based Systems*, vol. 10, no. 1, pp. 3–13, 1997.

[16] CARDOSO, A., VEALE, T., and WIGGINS, G. A., "Converging on the divergent: The history (and future) of the international joint workshops in computational creativity," *AI Magazine*, vol. 30, no. 3, p. 15, 2009.

[17] CARROLL, E. A., LATULIPE, C., FUNG, R., and TERRY, M., "Creativity factor evaluation: towards a standardized survey metric for creativity support," in *Proceedings of the seventh ACM conference on Creativity and cognition*, pp. 127–136, ACM, 2009.

[18] CHAMPANDARD, A. J., "Semantic style transfer and turning two-bit doodles into fine artworks," *arXiv preprint arXiv:1603.01768*, 2016.

[19] CHATFIELD, K., SIMONYAN, K., VEDALDI, A., and ZISSERMAN, A., "Return of the devil in the details: Delving deep into convolutional nets," *arXiv preprint arXiv:1405.3531*, 2014.

[20] CHEONG, Y. G., "A computational model of narrative generation for suspense," 2007.

[21] CHEONG, Y.-G. and YOUNG, R. M., "Narrative generation for suspense: Modeling and evaluation," in *Joint International Conference on Interactive Digital Storytelling*, pp. 144–155, Springer, 2008.

[22] COLTON, S., "Creativity versus the perception of creativity in computational systems.," in *AAAI Spring Symposium: Creative Intelligent Systems*, pp. 14–20, 2008.

[23] COLTON, S., GOODWIN, J., and VEALE, T., "Full face poetry generation," in *Proceedings of the Third International Conference on Computational Creativity*, pp. 95–102, 2012.

[24] COLTON, S., LOPEZ DE MANTARAS, R., STOCK, O., and OTHERS, "Computational creativity: Coming of age," 2009.

[25] COMPTON, K. and MATEAS, M., "Casual creators," in *Proceedings of the Sixth International Conference on Computational Creativity June*, vol. 228, 2015.

[26] COOK, M. and COLTON, S., "Ludus ex machina: Building a 3d game designer that competes alongside humans," in *Proceedings of the 5th international conference on computational creativity*, vol. 380, 2014.

[27] Csikszentmihalyi, M., *Creativity: Flow and the Psychology of Discovery.* HarperCollins, 2009.

[28] Cuffari, E. C., Di Paolo, E., and De Jaegher, H., "From participatory sense-making to language: there and back again," *Phenomenology and the Cognitive Sciences*, vol. 14, no. 4, pp. 1089–1125, 2015.

[29] Cuffari, E. C., Di Paolo, E., and De Jaegher, H., "From participatory sense-making to language: there and back again," *Phenomenology and the Cognitive Sciences*, vol. 14, no. 4, pp. 1089–1125, 2015.

[30] Davis, N., "Human-computer co-creativity: Blending human and computational creativity," in *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2013.

[31] Davis, N., Do, E. Y.-L., Gupta, P., and Gupta, S., "Computing harmony with perlogicart: perceptual logic inspired collaborative art," in *Proceedings of the 8th ACM conference on Creativity and cognition*, pp. 185–194, ACM, 2011.

[32] Davis, N., Hsiao, C.-P., Popova, Y., and Magerko, B., "An enactive model of creativity for computational collaboration and co-creation," in *Creativity in the Digital Age*, pp. 109–133, Springer, 2015.

[33] Davis, N., Hsiao, C.-P., Singh, K. Y., Li, L., Moningi, S., and Magerko, B., "Drawing apprentice: An enactive co-creative agent for artistic collaboration," in *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pp. 185–186, ACM, 2015.

[34] Davis, N., Hsiao, C.-P., Singh, K. Y., and Magerko, B., "Co-creative drawing agent with object recognition," in *Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2016.

[35] Davis, N., Hsiao, C.-P., Yashraj Singh, K., Li, L., and Magerko, B., "Empirically studying participatory sense-making in abstract drawing with a co-creative cognitive agent," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 196–207, ACM, 2016.

[36] Davis, N., Popova, Y., Sysoeven, I., Zhang, D., and Magerko, B., "Building artistic computer colleagues with an enactive model of creativity," in *International Conference on Computational Creativity, Ljubljuana Slovenia*, AAI, 2014.

[37] De Bono, E., *Lateral thinking for management: A handbook of creativity.* American Management Association, 1971.

[38] De Jaegher, H., "Social understanding through direct perception? yes, by interacting," *Consciousness and cognition*, vol. 18, no. 2, pp. 535–542, 2009.

[39] De Jaegher, H. and Di Paolo, E., "Participatory sense-making," *Phenomenology and the cognitive sciences*, vol. 6, no. 4, pp. 485–507, 2007.

[40] De Jaegher, H. and Di Paolo, E., "Making sense in participation: An enactive approach to social cognition," *Emerging communication*, vol. 10, p. 33, 2008.

[41] De Jaegher, H., Di Paolo, E., and Gallagher, S., "Can social interaction constitute social cognition?," *Trends in cognitive sciences*, vol. 14, no. 10, pp. 441–447, 2010.

[42] Di Paolo, E. A., Rohde, M., and Iizuka, H., "Sensitivity to social contingency or stability of interaction? modelling the dynamics of perceptual crossing," *New ideas in psychology*, vol. 26, no. 2, pp. 278–294, 2008.

[43] Dietrich, A., "The cognitive neuroscience of creativity," *Psychonomic bulletin & review*, vol. 11, no. 6, pp. 1011–1026, 2004.

[44] Dixon, D., Prasad, M., and Hammond, T., "icandraw: using sketch recognition and corrective feedback to assist a user in drawing human faces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 897–906, ACM, 2010.

[45] Do, E. Y.-L., "The right tool at the right time: investigation of freehand drawing as an interface to knowledge based design tools," 1999.

[46] Do, E. Y.-L., "Design sketches and sketch design tools," *Knowledge-Based Systems*, vol. 18, no. 8, pp. 383–405, 2005.

[47] Do, E. Y.-L. and Gross, M. D., "Environments for creativity: a lab for making things," in *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, pp. 27–36, ACM, 2007.

[48] Dorst, K. and Cross, N., "Creativity in the design process: co-evolution of problem–solution," *Design studies*, vol. 22, no. 5, pp. 425–437, 2001.

[49] Dwarakanath, S. and Blessing, L., "Design process ingredients: a comparison between group and individual work," *Analysing design activity*, pp. 93–116, 1994.

[50] Edmunds, P., "Swarmsketch," in *PDC*, pp. 21–22, 2006.

[51] Eitz, M., Hays, J., and Alexa, M., "How do humans sketch objects?," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 44–1, 2012.

[52] Engel, A. K., "Directive minds: how dynamics shapes cognition," *Enaction: Towards a New Paradigm for Cognitive Science*, pp. 219–243, 2010.

[53] Falkenhainer, B., Forbus, K. D., and Gentner, D., "The structure-mapping engine: Algorithm and examples," *Artificial intelligence*, vol. 41, no. 1, pp. 1–63, 1989.

[54] Fantasia, V., De Jaegher, H., and Fasulo, A., "We can work it out: an enactive look at cooperation," *Frontiers in psychology*, vol. 5, 2014.

[55] Fink, A., Benedek, M., Grabner, R. H., Staudt, B., and Neubauer, A. C., "Creativity meets neuroscience: Experimental tasks for the neuroscientific study of creative thinking," *Methods*, vol. 42, no. 1, pp. 68–76, 2007.

[56] Flagg, M. and Rehg, J. M., "Projector-guided painting," in *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pp. 235–244, ACM, 2006.

[57] Fleiss, J. L., "Measuring nominal scale agreement among many raters.," *Psychological bulletin*, vol. 76, no. 5, p. 378, 1971.

[58] French, R. M., "The computational modeling of analogy-making," *Trends in cognitive Sciences*, vol. 6, no. 5, pp. 200–205, 2002.

[59] Friston, K., "The free-energy principle: a rough guide to the brain?," *Trends in cognitive sciences*, vol. 13, no. 7, pp. 293–301, 2009.

[60] Friston, K., "The free-energy principle: a unified brain theory?," *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[61] Friston, K., Kilner, J., and Harrison, L., "A free energy principle for the brain," *Journal of Physiology-Paris*, vol. 100, no. 1, pp. 70–87, 2006.

[62] Froese, T., Iizuka, H., and Ikegami, T., "Embodied social interaction constitutes social cognition in pairs of humans: a minimalist virtual reality experiment," *Scientific reports*, vol. 4, 2014.

[63] Froese, T., McGann, M., Bigge, W., Spiers, A., and Seth, A. K., "The enactive torch: a new tool for the science of perception," *IEEE Transactions on Haptics*, vol. 5, no. 4, pp. 365–375, 2012.

[64] Froese, T. and Spiers, A., "Toward a phenomenological pragmatics of enactive perception," in *4th International Conference on Enactive Interfaces 2007*, p. 105, 2007.

[65] Froese, T. and Ziemke, T., "Enactive artificial intelligence: Investigating the systemic organization of life and mind," *Artificial Intelligence*, vol. 173, no. 3, pp. 466–500, 2009.

[66] Fuchs, T. and De Jaegher, H., "Enactive intersubjectivity: Participatory sense-making and mutual incorporation," *Phenomenology and the Cognitive Sciences*, vol. 8, no. 4, pp. 465–486, 2009.

[67] Gabora, L., "Revenge of the neurds: characterizing creative thought in terms of the structure and dynamics of memory," *Creativity Research Journal*, vol. 22, no. 1, pp. 1–13, 2010.

[68] Gaspar, J. M. and McDonald, J. J., "Suppression of salient objects prevents distraction in visual search," *The Journal of Neuroscience*, vol. 34, no. 16, pp. 5658–5666, 2014.

[69] Gatys, L. A., Ecker, A. S., and Bethge, M., "A neural algorithm of artistic style," *arXiv preprint arXiv:1508.06576*, 2015.

[70] Gatys, L. A., Ecker, A. S., and Bethge, M., "Image style transfer using convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2414–2423, 2016.

[71] Gero, J. S., "Computational models of creative designing based on situated cognition," in *Proceedings of the 4th conference on Creativity & cognition*, pp. 3–10, ACM, 2002.

[72] Gibson, J. J., *The ecological approach to visual perception.* Routledge, 1986.

[73] Gick, M. L. and Holyoak, K. J., "Analogical problem solving," *Cognitive psychology*, vol. 12, no. 3, pp. 306–355, 1980.

[74] Glaser, B. G. and Strauss, A. L., *The discovery of grounded theory: Strategies for qualitative research.* Transaction Books, 2009.

[75] Goldschmidt, G., "The dialectics of sketching," *Creativity research journal*, vol. 4, no. 2, pp. 123–143, 1991.

[76] Gow, J. and Corneli, J., "Towards generating novel games using conceptual blending," in *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*, 2015.

[77] Grace, K. and Maher, M. L., "What to expect when youre expecting: the role of unexpectedness in computationally evaluating creativity," in *Proceedings of the Fifth International Conference on Computational Creativity (ICCC'14), Ljubljana, Slovenia*, 2014.

[78] Gregor, K., Danihelka, I., Graves, A., Rezende, D. J., and Wierstra, D., "Draw: A recurrent neural network for image generation," *arXiv preprint arXiv:1502.04623*, 2015.

[79] Guilford, J. P., "Creativity: Retrospect and prospect*," *The Journal of Creative Behavior*, vol. 4, no. 3, pp. 149–168, 1970.

[80] Gulz, A. and Haake, M., "Design of animated pedagogical agentsa look at their look," *International Journal of Human-Computer Studies*, vol. 64, no. 4, pp. 322–339, 2006.

[81] Havasi, C., Speer, R., and Alonso, J., "Conceptnet 3: a flexible, multilingual semantic network for common sense knowledge," in *Recent advances in natural language processing*, pp. 27–29, Citeseer, 2007.

[82] Havelange, V., "The ontological constitution of cognition and the epistemological constitution of cognitive science: phenomenology, enaction and technology," *Enaction: Toward a new paradigm for cognitive science*, pp. 335–359, 2010.

[83] Helms, M. and Goel, A., "Analogical problem evolution in biologically inspired design," *Design Computing and Cognition, June*, vol. 9, 2012.

[84] Hewett, T., Czerwinski, M., Terry, M., Nunamaker, J., Candy, L., Kules, B., and Sylvan, E., "Creativity support tool evaluation methods and metrics," *Creativity Support Tools*, pp. 10–24, 2005.

[85] Hodhod, R., Piplica, A., and Magerko, B., "A formal model of shared mental models for computational improvisational agents," in *Proceedings of the 12th Annual Conference on Intelligent Virtual Agents*, 2012.

[86] Hodhod, R. and Magerko, B., "Pharaoh: Conceptual blending of cognitive scripts for computationally creative agents," in *The Twenty-Seventh International Flairs Conference*, 2014.

[87] Hodhod, R. and Magerko, B., "Closing the cognitive gap between humans and interactive narrative agents using shared mental models," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 135–146, ACM, 2016.

[88] Hodhod, R., Magerko, B., and Gawish, M., "Pharaoh: Context-based structural retrieval of cognitive scripts," *International Journal of Information Retrieval Research (IJIRR)*, vol. 2, no. 3, pp. 58–71, 2012.

[89] Hoffman, G. and Weinberg, G., "Gesture-based human-robot jazz improvisation," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pp. 582–587, IEEE, 2010.

[90] Holyoak, K. J. and Thagard, P., "Analogical mapping by constraint satisfaction," *Cognitive science*, vol. 13, no. 3, pp. 295–355, 1989.

[91] Hsiao, C.-P., "Solidsketch: Toward enactive interactions for semantic model creation," in *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pp. 329–330, ACM, 2015.

[92] Hsiao, C.-P., Davis, N., and Do, E. Y.-L., "Dancing on the desktop - gesture modeling system to augment design cognition," in *ACADIA*, pp. 419–428, 2012.

[93] Huizinga, J., *Homo Ludens: A Study of the Play-element in Cult*. Beacon Press, 1955.

[94] Jacob, M. and Magerko, B., "Viewpoints ai," in *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pp. 361–362, ACM, 2015.

[95] Jacob, M., Zook, A., and Magerko, B., "Viewpoints ai: Procedurally representing and reasoning about gestures," in *Proceedings of DiGRA*, 2013.

[96] Jeannette, Y., Helms, M., Vattam, S., and Goel, A. K., "Evaluating biological systems for their potential in engineering design," *Advances in Natural Science*, vol. 3, no. 2, pp. 27–40, 2010.

[97] Jordan, B. and Henderson, A., "Interaction analysis: Foundations and practice," *The journal of the learning sciences*, vol. 4, no. 1, pp. 39–103, 1995.

[98] Kingma, D. and Ba, J., "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[99] Koestler, A., "The act of creation," 1964.

[100] Lakoff, G., "Cognitive models and prototype theory," *Concepts: Core Readings*, pp. 391–421, 1999.

[101] Lee, Y. J., Zitnick, C. L., and Cohen, M. F., "Shadowdraw: real-time user guidance for freehand drawing," in *ACM Transactions on Graphics (TOG)*, vol. 30, p. 27, ACM, 2011.

[102] Li, B., Zook, A., Davis, N., and Riedl, M. O., "Goal-driven conceptual blending: A computational approach for creativity," in *Proc. 2012 International Conference on Computational Creativity, Dublin, Ireland*, 2012.

[103] Lubart, T., "How can computers be partners in the creative process: classification and commentary on the special issue," *International Journal of Human-Computer Studies*, vol. 63, no. 4, pp. 365–369, 2005.

[104] Mace, M.-A. and Ward, T., "Modeling the creative process: A grounded theory analysis of creativity in the domain of art making," *Creativity Research Journal*, vol. 14, no. 2, pp. 179–192, 2002.

[105] Magerko, B., Fiesler, C., Baumer, A., and Fuller, D., "Bottoms up: improvisational micro-agents," in *Proceedings of the Intelligent Narrative Technologies III Workshop*, p. 8, ACM, 2010.

[106] Mahadevan, A., Freeman, J., Magerko, B., and Martinez, J. C., "Earsketch: Teaching computational music remixing in an online web audio based learning environment," in *Web Audio Conference*, 2015.

[107] Maher, M. L., "Creativity, computation, and interaction," in *Proc of CHI*, vol. 2008, 2008.

[108] Maher, M. L., Bilda, Z., and Gül, L. F., "Impact of collaborative virtual environments on design behaviour," in *Design computing and cognition06*, pp. 305–321, Springer, 2006.

[109] Maher, M. L., Brady, K., and Fisher, D. H., "Computational models of surprise in evaluating creative design," in *Proceedings of the fourth international conference on computational creativity*, vol. 147, Citeseer, 2013.

[110] Mamykina, L., Candy, L., and Edmonds, E., "Collaborative creativity," *Communications of the ACM*, vol. 45, no. 10, pp. 96–99, 2002.

[111] Mayer, R. E., "22 fifty years of creativity research," *Handbook of creativity*, p. 449, 1999.

[112] McCoid, S., Freeman, J., Magerko, B., Michaud, C., Jenkins, T., Mcklin, T., and Kan, H., "Earsketch: An integrated approach to teaching introductory computer music," *Organised Sound*, vol. 18, no. 02, pp. 146–160, 2013.

[113] McCorduck, P., *Aaron's code*. WH Freeman & Co., 1991.

[114] Merleau-Ponty, M., *Phenomenology of perception*. Motilal Banarsidass Publishe, 1996.

[115] Murdock, M. C. and Puccio, G. J., "A contextual organizer for conducting creativity research," *Understanding and recognizing creativity: The emergence of a discipline*, pp. 249–280, 1993.

[116] Nakakoji, K., "Meanings of tools, support, and uses for creative design processes," in *Proc. International Design Research Symposium*, pp. 156–165, 2006.

[117] Nelligan, T., Polsley, S., Ray, J., Helms, M., Linsey, J., and Hammond, T., "Mechanix: a sketch-based educational interface," in *Proceedings of the 20th International Conference on Intelligent User Interfaces Companion*, pp. 53–56, ACM, 2015.

[118] Nersessian, N., *Creating scientific concepts*. The MIT Press, 2008.

[119] Nersessian, N. J., *Faraday to Einstein: Constructing meaning in scientific theories*. Springer, 1984.

[120] Nersessian, N. J., "Model-based reasoning in conceptual change," in *Model-based reasoning in scientific discovery*, pp. 5–22, Springer, 1999.

[121] Nersessian, N. J., "The cognitive basis of model-based reasoning in science," *The cognitive basis of science*, pp. 133–153, 2002.

[122] Netzer, Y., Gabay, D., Goldberg, Y., and Elhadad, M., "Gaiku: Generating haiku with word associations norms," in *Proceedings of the Workshop on Computational Approaches to Linguistic Creativity*, pp. 32–39, Association for Computational Linguistics, 2009.

[123] Newell, A., Shaw, J. C., and Simon, H. A., *The processes of creative thinking*. Rand Corporation, 1959.

[124] Noë, A., *Action in perception*. MIT Press, 2004.

[125] Noë, A., "Vision without representation," *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems*, pp. 245–256, 2010.

[126] Norman, D. A., "Affordance, conventions, and design," *interactions*, vol. 6, no. 3, pp. 38–43, 1999.

[127] Oliveira, H., "Automatic generation of poetry: an overview," *Universidade de Coimbra*, 2009.

[128] Oullier, O. and Basso, F., "Embodied economics: how bodily information shapes the social coordination dynamics of decision-making," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 365, no. 1538, pp. 291–301, 2010.

[129] ONeill, B., *A computational model of suspense for the augmentation of intelligent story generation*. PhD thesis, Georgia Institute of Technology, 2013.

[130] Pease, A., Winterstein, D., and Colton, S., "Evaluating machine creativity," in *Workshop on Creative Systems, 4th International Conference on Case Based Reasoning*, pp. 129–137, 2001.

[131] Peinado, F. and Gervás, P., "Evaluation of automatic generation of basic stories," *New Generation Computing*, vol. 24, no. 3, pp. 289–302, 2006.

[132] Pellegrini, A. D. and Smith, P. K., "Physical activity play: The nature and function of a neglected aspect of play," *Child development*, vol. 69, no. 3, pp. 577–598, 1998.

[133] Philipona, D., O'Regan, J. K., and Nadal, J.-P., "Is there something out there? inferring space from sensorimotor dependencies," *Neural computation*, vol. 15, no. 9, pp. 2029–2049, 2003.

[134] Poon, J. and Maher, M. L., "Co-evolution and emergence in design," *Artificial Intelligence in Engineering*, vol. 11, no. 3, pp. 319–327, 1997.

[135] Popova, Y. B., "Narrativity and enaction: the social nature of literary narrative understanding," *Frontiers in psychology*, vol. 5, 2014.

[136] Prinz, J. J. and Barsalou, L. W., "Acquisition and productivity in perceptual symbol systems: An account of mundane creativity," *Creativity, Cognition, and Knowledge. Ed Terry Dartnall. Westport, CT: Praeger*, pp. 105–38, 2002.

[137] Rashel, F. and Manurung, R., "Pemuisi: a constraint satisfaction-based generator of topical indonesian poetry," in *5th international conference on computational creativity, ICCC*, 2014.

[138] Rawlinson, J. G., *Creative thinking and brainstorming*. Gower Farnborough, Hants, 1981.

[139] Robertson, J. and Young, R. M., "Automated gameplay generation from declarative world representations," in *Proceedings of the 11th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment,(to appear)*, 2015.

[140] Rosch, E., "Cognitive representations of semantic categories.," *Journal of experimental psychology: General*, vol. 104, no. 3, p. 192, 1975.

[141] Rucinska, Z. and Reijmers, E., "Enactive account of pretend play and its application to therapy," *Frontiers in psychology*, vol. 6, 2015.

[142] Runco, M. A. and Richards, R., *Eminent creativity, everyday creativity, and health*. Greenwood Publishing Group, 1997.

[143] Sandini, G., Metta, G., and Vernon, D., "Robotcub: An open framework for research in embodied cognition," in *Humanoid Robots, 2004 4th IEEE/RAS International Conference on*, vol. 1, pp. 13–32, IEEE, 2004.

[144] Sawyer, K., "The cognitive neuroscience of creativity: a critical review," *Creativity Research Journal*, vol. 23, no. 2, pp. 137–154, 2011.

[145] Sawyer, R. K., *Explaining creativity: The science of human innovation*. Oxford University Press, 2011.

[146] Sawyer, R. K. and DeZutter, S., "Distributed creativity: How collective creations emerge from collaboration.," *Psychology of Aesthetics, Creativity, and the Arts*, vol. 3, no. 2, p. 81, 2009.

[147] Sawyer, R. K., *Pretend play as improvisation: Conversation in the preschool classroom*. Psychology Press, 1997.

[148] Schön, D. A., "Designing as reflective conversation with the materials of a design situation," *Knowledge-Based Systems*, vol. 5, no. 1, pp. 3–14, 1992.

[149] Schorlemmer, M., Smaill, A., Kühnberger, K.-U., Kutz, O., Colton, S., Cambouropoulos, E., and Pease, A., "Coinvent: Towards a computational concept invention theory," 2014.

[150] Schwartenbeck, P., FitzGerald, T., Dolan, R., and Friston, K., "Exploration, novelty, surprise, and free energy minimization," *Frontiers in psychology*, vol. 4, p. 710, 2013.

[151] Segal, E., "Incubation in insight problem solving," *Creativity Research Journal*, vol. 16, no. 1, pp. 141–148, 2004.

[152] Shneiderman, B., "Creativity support tools," *Communications of the ACM*, vol. 45, no. 10, pp. 116–120, 2002.

[153] Shneiderman, B., "Creativity support tools: accelerating discovery and innovation," *Communications of the ACM*, vol. 50, no. 12, pp. 20–32, 2007.

[154] Shneiderman, B., Fischer, G., Czerwinski, M., Resnick, M., Myers, B., Candy, L., Edmonds, E., Eisenberg, M., Giaccardi, E., Hewett, T., and others, "Creativity support tools: Report from a us national science foundation sponsored workshop," *International Journal of Human-Computer Interaction*, vol. 20, no. 2, pp. 61–77, 2006.

[155] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[156] Smith, S. M., Ward, T. B., and Finke, R. A., *The Creative Cognition Approach: Edited by Steven M. Smith, Thomas B. Ward, and Ronald A. Finke.* The MIT Press, 1995.

[157] Sowden, P. T., Pringle, A., and Gabora, L., "The shifting sands of creative thinking: Connections to dual-process theory," *Thinking & Reasoning*, vol. 21, no. 1, pp. 40–60, 2015.

[158] Srinivasan, N., "Cognitive neuroscience of creativity: Eeg based approaches," *Methods*, vol. 42, no. 1, pp. 109–116, 2007.

[159] Steffensen, S. and Pedersen, S. B., "Temporal dynamics in human interaction," *Cybernetics & Human Knowing*, vol. 21, no. 1-2, pp. 80–97, 2014.

[160] Steffensen, S. V., "Care and conversing in dialogical systems," *Language Sciences*, vol. 34, no. 5, pp. 513–531, 2012.

[161] Steiner, P. and Stewart, J., "From autonomy to heteronomy (and back): The enaction of social life," *Phenomenology and the Cognitive Sciences*, vol. 8, no. 4, pp. 527–550, 2009.

[162] Stewart, J. R., Gapenne, O., and Di Paolo, E. A., *Enaction: Toward a new paradigm for cognitive science.* MIT Press, 2010.

[163] Suchman, L. A., *Plans and situated actions: the problem of human-machine communication.* Cambridge university press, 1987.

[164] Suchman, L., "Response to vera and simon's situated action: A symbolic interpretation," *Cognitive Science*, vol. 17, no. 1, pp. 71–75, 1993.

[165] Suchman, L., *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press, 2007.

[166] Sutton-Smith, B., "The role of toys in the instigation of playful creativity," *Creativity Research Journal*, vol. 5, no. 1, pp. 3–11, 1992.

[167] Suwa, M. and Tversky, B., "What do architects and students perceive in their design sketches? a protocol analysis," *Design studies*, vol. 18, no. 4, pp. 385–403, 1997.

[168] Torrance, E., "Predictive validity of the torrance tests of creative thinking*," *The Journal of creative behavior*, vol. 6, no. 4, pp. 236–262, 1972.

[169] Van der Lugt, R., "How sketching can affect the idea generation process in design group meetings," *Design studies*, vol. 26, no. 2, pp. 101–122, 2005.

[170] Varela, F. J., Thompson, E. T., and Rosch, E., *The embodied mind: Cognitive science and human experience*. The MIT Press, 1991.

[171] Vattam, S., Wiltgen, B., Helms, M., Goel, A. K., and Yen, J., "Dane: fostering creativity in and through biologically inspired design," in *Design Creativity 2010*, pp. 115–122, Springer, 2011.

[172] Veale, T., "Coming good and breaking bad: Generating transformative character arcs for use in compelling stories," in *Proceedings of ICCC-2014, the 5th International Conference on Computational Creativity, Ljubljana, June 2014*, 2014.

[173] Veale, T. and Hao, Y., "A fluid knowledge representation for understanding and generating creative metaphors," in *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pp. 945–952, Association for Computational Linguistics, 2008.

[174] Vernon, D., "Enaction as a conceptual framework for developmental cognitive robotics," *Paladyn, Journal of Behavioral Robotics*, vol. 1, no. 2, pp. 89–98, 2010.

[175] Vernon, D., *Artificial cognitive systems: A primer*. MIT Press, 2014.

[176] Vernon, D., Metta, G., and Sandini, G., "A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 2, p. 151, 2007.

[177] Wardrip-Fruin, N., *Expressive Processing: Digital fictions, computer games, and software studies*. MIT press, 2009.

[178] Weizenbaum, J., "Elizaa computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.

[179] Wiggins, G. A., "A preliminary framework for description, analysis and comparison of creative systems," *Knowledge-Based Systems*, vol. 19, no. 7, pp. 449–458, 2006.

[180] Wiggins, G. A., "Searching for computational creativity," *New Generation Computing*, vol. 24, no. 3, pp. 209–222, 2006.

[181] Yamamoto, Y. and Nakakoji, K., "Interaction design of tools for fostering creativity in the early stages of information design," *International Journal of Human-Computer Studies*, vol. 63, no. 4, pp. 513–535, 2005.

[182] Yaner, P. W. and Goel, A. K., "Visual analogy: Viewing analogical retrieval and mapping as constraint satisfaction problems," *Applied Intelligence*, vol. 25, no. 1, pp. 91–105, 2006.

[183] Yannakakis, G. N., Liapis, A., and Alexopoulos, C., "Mixed-initiative co-creativity," in *Proceedings of the 9th Conference on the Foundations of Digital Games*, 2014.

[184] Yen, J., Weissburg, M., Helms, M., and Goel, A., "Biologically inspired design: a tool for interdisciplinary education," 2011.

[185] Yokochi, S. and Okada, T., "Creative cognitive process of art making: A field study of a traditional chinese ink painter," *Creativity Research Journal*, vol. 17, no. 2-3, pp. 241–255, 2005.

[186] Yonge, G. D., "Structure of experience and functional fixedness.," *Journal of Educational Psychology*, vol. 57, no. 2, p. 115, 1966.

[187] Yu, Q., Yang, Y., Song, Y.-Z., Xiang, T., and Hospedales, T., "Sketch-a-net that beats humans," *arXiv preprint arXiv:1501.07873*, 2015.

[188] Zook, A., Magerko, B., and Riedl, M., "Formally modeling pretend object play," in *Proceedings of the 8th ACM Conference on Creativity and Cognition*, pp. 147–156, ACM, 2011.

[189] Zook, A. E., Riedl, M. O., and Magerko, B. S., "Understanding human creativity for computational play," in *8th ACM Conference on Creativity and Cognition*, 2011.